

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2006, 18:53:48 ; Search time 4.24897 Seconds  
(without alignments)  
1034.083 Million cell updates/sec

Title: US-10-525-567-1  
Perfect score: 49  
Sequence: 1 AAVNHQKSA 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_21.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*  
9: Geneseqp2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	49	100.0	10	8	ADL27604 Honey bee
2	49	100.0	524	8	ADL27606 Honey bee
3	49	100.0	544	8	ADL27618 Honey bee
4	49	100.0	544	8	ADL27616 Honey bee
5	36	73.5	347	4	ABb64181 Drosophil
6	36	73.5	503	4	ABb63773 Drosophil
7	36	73.5	599	4	ABg14172 Novel hum
8	36	73.5	599	4	AAU32811 Novel hum
9	35	71.4	535	6	AAU40932 Propionib
10	35	71.4	535	6	ABm37451 Propionib
11	35	71.4	664	7	ADD31051 Plant yie
12	35	71.4	664	8	ADl44067 Plant tra
13	35	71.4	5509	8	ADN96832 Bugula br
14	35	71.4	6940	8	ADN96829 Bugula br
15	34	69.4	284	8	ADS30840 Bacterial
16	34	69.4	291	4	ABg07730 Novel hum
17	34	69.4	300	4	ABg14415 Novel hum
18	34	69.4	307	6	ABm67540 Photorhab
19	34	69.4	387	5	ABp69562 Human pol
20	34	69.4	445	4	ABg05552 Novel hum
21	34	69.4	453	4	ABg14171 Novel hum
22	34	69.4	454	4	ABg06133 Novel hum
23	34	69.4	479	2	AAR99230 Barley xy
24	34	69.4	562	8	ADQ67724 Novel hum

25	34	69.4	604	9	AEA20216	Aea20216 Novel hum
26	33	67.3	71	4	AAU64987	Aau64987 Propionib
27	33	67.3	71	6	ABM61506	Abm61506 Propionib
28	33	67.3	250	6	ABU20971	Abu20971 Protein e
29	33	67.3	291	4	ABG08383	Abg08383 Novel hum
30	33	67.3	375	4	ABB68042	Abb68042 Drosophil
31	33	67.3	380	4	AAU32883	Aau32883 Novel hum
32	33	67.3	809	6	ABP77288	Abp77288 N. gonorr
33	33	67.3	810	9	ABE91535	Abeg91535 Microbial
34	33	67.3	1103	8	ADQ67756	Adq67756 Novel hum
35	33	67.3	3572	5	ABG95659	Abg95659 Human nuc
36	32	65.3	53	4	AAU09275	Aau09275 Human pol
37	32	65.3	66	4	AAU66448	Aau66448 Propionib
38	32	65.3	66	6	ABM62967	Abm62967 Propionib
39	32	65.3	71	4	ABG03124	Abg03124 Novel hum
40	32	65.3	77	8	ADK48037	Adk48037 Streptoco
41	32	65.3	79	4	AAU54537	Aau54537 Propionib
42	32	65.3	79	6	ABM51056	Abm51056 Propionib
43	32	65.3	83	4	ABG26402	Abg26402 Novel hum
44	32	65.3	242	9	ABM94288	Abm94288 M. xanthu
45	32	65.3	249	7	ADC97121	Adc97121 E. faeciu

## ALIGNMENTS

RESULT 1  
ADL27604  
ID ADL27604 standard; peptide; 10 AA.  
XX  
AC ADL27604;  
XX  
XX  
DT 03-JUN-2004 (first entry)  
XX  
DE Honey bee RJP70 N-terminal peptide, SEQ ID 1.  
XX  
KW Antiallergic; Antiasthmatic; Antiinflammatory; Dermatological; honey bee;  
KW royal jelly protein; allergy; pollinosis; atopic dermatitis;  
KW contact hypersensitivity; bronchial asthma; allergic rhinitis; RJP70.  
XX  
OS Apis mellifera.  
XX  
XX WO2004019971-A1.  
PN  
PD 11-MAR-2004.  
XX  
XX 26-AUG-2003; 2003WO-JP010795.  
PF  
PR 29-AUG-2002; 2002JP-00252087.  
PR 30-JAN-2003; 2003JP-00022776.  
XX  
XX (HAYA/) HAYASHIBARA K.  
XX  
XX Okamoto I, Arai N, Kohno K, Kurimoto M, Sano O;  
XX  
XX WPI; 2004-248191/23.  
XX  
XX Antiallergic agents for relieving symptoms accompanying e.g. pollinosis,  
XX atopic dermatitis, contact hypersensitivity, bronchial asthma and  
XX allergic rhinitis, comprise protein originated from royal jelly.  
XX  
XX Claim 1; SEQ ID NO 1; 78pp; Japanese.  
XX  
XX The present invention relates to novel antiallergic agents, which  
XX comprise as an active ingredient, peptides ADL27604 and ADL27605, which  
XX are N-terminal peptides of honey bee royal jelly proteins (RJP, ADL27606  
XX and ADL27607). The agents can be used to relieve symptoms accompanying an  
XX allergic disease e.g. pollinosis, atopic dermatitis, contact  
XX hypersensitivity, bronchial asthma and allergic rhinitis, which are  
XX applicable in foods, cosmetics and drugs.  
XX  
XX Sequence 10 AA;

```

Query Match      100.0%; Score 49; DB 8; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0051;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAVNHQRKSA 10
Db 1 AAVNHQRKSA 10
|||||

RESULT 2
ADL27606
ID ADL27606 standard; protein; 524 AA.
XX AC ADL27606;
XX XX
XX 03-JUN-2004 (first entry)
XX DE Honey bee RJP70 mature protein, SEQ ID 3.
XX KW Antiallergic; Antiasthmatic; Antiinflammatory; Dermatological; honey bee;
XX KW royal jelly protein; allergy; pollinosis; atopic dermatitis;
XX KW contact hypersensitivity; bronchial asthma; allergic rhinitis; RJP70.
XX OS Apis mellifera.
XX PN WO2004019971-A1.
XX XX
XX 11-MAR-2004.
XX PF 26-AUG-2003; 2003WO-JP010795.
XX PR 29-AUG-2002; 2002JP-00252087.
XX PR 30-JAN-2003; 2003JP-00022776.
XX XX
XX PA (HAYA/) HAYASHIBARA K.
XX XX
XX OKamoto I, Arai N, Kohno K, Kurimoto M, Sano O;
XX WPI; 2004-248191/23.
XX DR N-PSDB; ADL27608.
XX XX
XX Antiallergic agents for relieving symptoms accompanying e.g. pollinosis,
XX PT atopic dermatitis, contact hypersensitivity, bronchial asthma and
XX PT allergic rhinitis, comprise protein originated from royal jelly.
XX XX
XX Claim 1; SEQ ID NO 3; 78pp; Japanese.
XX PS
XX CC The present invention relates to novel antiallergic agents, which
XX CC comprise as an active ingredient, peptides ADL27604 and ADL27605, which
XX CC are N-terminal peptides of honey bee royal jelly proteins (RJP, ADL27606
XX CC and ADL27607). The agents can be used to relieve symptoms accompanying an
XX CC allergic disease e.g. pollinosis, atopic dermatitis, contact
XX CC hypersensitivity, bronchial asthma and allergic rhinitis, which are
XX CC applicable in foods, cosmetics and drugs.
XX XX
XX Sequence 524 AA;
XX QY
XX Db

Query Match      100.0%; Score 49; DB 8; Length 524;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAVNHQRKSA 10
Db 1 AAVNHQRKSA 10
|||||

RESULT 3
ADL27618
ID ADL27618 standard; protein; 544 AA.
XX AC ADL27618;
XX XX
XX 03-JUN-2004 (first entry)
XX DT

Query Match      100.0%; Score 49; DB 8; Length 544;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAVNHQRKSA 10
Db 21 AAVNHQRKSA 30
|||||

RESULT 4
ADL27616
ID ADL27616 standard; protein; 544 AA.
XX AC ADL27616;
XX XX
XX 03-JUN-2004 (first entry)
XX DE Honey bee RJP70.
XX KW Antiallergic; Antiasthmatic; Antiinflammatory; Dermatological; honey bee;
XX KW royal jelly protein; allergy; pollinosis; atopic dermatitis;
XX KW contact hypersensitivity; bronchial asthma; allergic rhinitis; RJP70.
XX OS Apis mellifera.
XX XX

```

PH Key Location/Qualifiers  
 FT Peptide 1..20  
 FT /label= Signal\_peptide  
 FT Protein 21..544  
 FT /label= Mature\_protein  
 XX WO2004019971-A1.  
 XX 11-MAR-2004.  
 XX 26-AUG-2003; 2003WO-JP010795.  
 XX 29-AUG-2002; 2002JP-00252087.  
 PR 30-JAN-2003; 2003JP-00022776.  
 XX (HAYA/) HAYASHIBARA K.  
 XX Okamoto I, Arai N, Kohno K, Kurimoto M, Sano O;  
 XX WPI; 2004-248191/23.  
 DR N-PSDB; ADL27608.  
 XX Antiallergic agents for relieving symptoms accompanying e.g. pollinosis,  
 PT atopic dermatitis, contact hypersensitivity, bronchial asthma and  
 PT allergic rhinitis, comprise protein originated from royal jelly.  
 XX Disclosure; Page 59-62; 78pp; Japanese.  
 XX The present invention relates to novel antiallergic agents, which  
 CC comprise as an active ingredient, peptides ADL27604 and ADL27605, which  
 CC are N-terminal peptides of honey bee royal jelly proteins (RJP, ADL27606  
 CC and ADL27607). The agents can be used to relieve symptoms accompanying an  
 CC allergic disease e.g. pollinosis, atopic dermatitis, contact  
 CC hypersensitivity, bronchial asthma and allergic rhinitis, which are  
 CC applicable in foods, cosmetics and drugs. The present sequence is a honey  
 CC bee RJP.  
 XX Sequence 544 AA;  
 SQ Query Match 100.0%; Score 49; DB 8; Length 544;  
 Best Local Similarity 100.0%; Pred. No. 0.35;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 AAVNHQRKSA 10  
 Db |||||  
 21 AAVNHQRKSA 30  
 RESULT 5  
 ABB64181  
 ID ABB64181 standard; protein; 347 AA.  
 XX AC ABB64181;  
 XX 26-MAR-2002 (first entry)  
 DT Drosophila melanogaster polypeptide SEQ ID NO 19335.  
 DE Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 KW Drosophila melanogaster.  
 OS WO200171042-A2.  
 PN 27-SEP-2001.  
 XX 23-MAR-2001; 2001WO-US009231.  
 PF 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX (PEKE ) PE CORP NY.  
 PA Venter JC, Adams M, Li PWD, Myers EW;  
 PI WPI; 2001-656860/75.  
 DR N-PSDB; ABL07876.  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.  
 PT Disclosure; SEQ ID NO 18111; 21pp + Sequence Listing; English.  
 PS The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ffp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 347 AA;  
 SQ Query Match 73.5%; Score 36; DB 4; Length 347;  
 Best Local Similarity 77.8%; Pred. No. 89;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 AAVNHQRKS 9  
 Db |||||  
 145 AAVNHQRQS 153  
 RESULT 6  
 ABB63773  
 ID ABB63773 standard; protein; 503 AA.  
 XX AC ABB63773;  
 XX 26-MAR-2002 (first entry)  
 DT Drosophila melanogaster polypeptide SEQ ID NO 1811.  
 DE Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 KW Drosophila melanogaster.  
 OS WO200171042-A2.  
 PN 27-SEP-2001.  
 XX 23-MAR-2001; 2001WO-US009231.  
 PF 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX (PEKE ) PE CORP NY.  
 PA Venter JC, Adams M, Li PWD, Myers EW;  
 PI WPI; 2001-656860/75.  
 DR N-PSDB; ABL07876.  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.  
 PT Disclosure; SEQ ID NO 18111; 21pp + Sequence Listing; English.  
 PS The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is

XX Venter JC, Adams M, Li PWD, Myers EW;  
 PI WPI; 2001-656860/75.  
 DR N-PSDB; ABL08284.  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.  
 PT Disclosure; SEQ ID NO 19335; 21pp + Sequence Listing; English.  
 PS The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ffp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 347 AA;  
 SQ Query Match 73.5%; Score 36; DB 4; Length 347;  
 Best Local Similarity 77.8%; Pred. No. 89;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 AAVNHQRKS 9  
 Db |||||  
 145 AAVNHQRQS 153  
 RESULT 6  
 ABB63773  
 ID ABB63773 standard; protein; 503 AA.  
 XX AC ABB63773;  
 XX 26-MAR-2002 (first entry)  
 DT Drosophila melanogaster polypeptide SEQ ID NO 1811.  
 DE Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 KW Drosophila melanogaster.  
 OS WO200171042-A2.  
 PN 27-SEP-2001.  
 XX 23-MAR-2001; 2001WO-US009231.  
 PF 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX (PEKE ) PE CORP NY.  
 PA Venter JC, Adams M, Li PWD, Myers EW;  
 PI WPI; 2001-656860/75.  
 DR N-PSDB; ABL07876.  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.  
 PT Disclosure; SEQ ID NO 18111; 21pp + Sequence Listing; English.  
 PS The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 503 AA;  
 Query Match 73.5%; Score 36; DB 4; Length 503;  
 Best Local Similarity 77.8%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 AVNHQKSA 10  
 Db 29 AVRHQRSA 37  
 |||||  
 |||||  
 RESULT 7  
 ABL14172  
 ID ABL14172 standard; protein; 599 AA.  
 XX  
 AC ABL14172;  
 XX  
 DT 18-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #14163.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US008631.  
 XX  
 PR 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Drmanac RT, Liu C, Tang YT;  
 XX  
 DR WPI; 2001-639362/73.  
 DR N-PSDB; AAS78359.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX  
 PS Claim 20; SEQ ID NO 44531; 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABC0010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 599 AA;  
 Query Match 73.5%; Score 36; DB 4; Length 599;  
 Best Local Similarity 75.0%; Pred. No. 1.6e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AAVNHQK 8  
 Db 141 AAINHQK 148  
 |||||  
 |||||  
 RESULT 8  
 AAU32811  
 ID AAU32811 standard; protein; 599 AA.  
 XX  
 AC AAU32811;  
 XX  
 DT 18-DEC-2001 (first entry)  
 XX  
 DE Novel human secreted protein #3302.  
 XX  
 KW Human; vaccination; gene therapy; nutritional supplement;  
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200179449-A2.  
 XX  
 PD 25-OCT-2001.  
 XX  
 PF 16-APR-2001; 2001WO-US008656.  
 XX  
 PR 18-APR-2000; 2000US-00552929.  
 PR 26-JAN-2001; 2001US-00770160.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Drmanac RT;  
 XX  
 DR WPI; 2001-611725/70.  
 XX  
 PT Nucleic acids encoding a range of human polypeptides, useful in genetic  
 PT vaccination, testing and therapy.  
 XX  
 PS Claim 20; Page 668-669; 765pp; English.  
 XX  
 CC The invention relates to novel human secreted polypeptides. The  
 CC polypeptides and antibodies to the polypeptides are useful for  
 CC determining the presence of or predisposition to a disease associated  
 CC with altered levels of polypeptide. The polypeptides are also useful for  
 CC identifying agents (agonists and antagonists) that bind to them. Cells  
 CC expressing the proteins are useful for identifying a therapeutic agent  
 CC for use in treatment of a pathology related to aberrant expression or  
 CC physiological interactions of the polypeptide. Vectors comprising the  
 CC nucleic acids encoding the polypeptides and cells genetically engineered  
 CC to express them are also useful for producing the proteins. The proteins  
 CC are useful in genetic vaccination, testing and therapy, and can be used  
 CC as nutritional supplements. They may be used to increase stem cell  
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
 CC and/or nerve tissue growth or regeneration; immune suppression and/or  
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
 CC AAU29510-AAU33304 represent the amino acid sequences of novel human  
 CC secreted proteins of the invention  
 XX  
 SQ Sequence 599 AA;



Query Match 73.5%; Score 36; DB 4; Length 599;  
Best Local Similarity 75.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAVNHQK 8  
Db 141 AAVNHQK 148

RESULT 9  
AAU40932  
ID AAU40932 standard; protein; 535 AA.  
XX AC AAU40932;  
XX DT 13-FEB-2002 (first entry)  
XX DE Propionibacterium acnes immunogenic protein #1828.  
XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
XX KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
XX KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
XX KW dermatological; osteopathic; neuroprotectant.

OS Propionibacterium acnes.  
XX WO200181581-A2.  
XX PN 01-NOV-2001.  
XX PD 20-APR-2001; 2001WO-US012865.  
XX PF 21-APR-2000; 2000US-0199047P.  
XX PR 02-JUN-2000; 2000US-0208841P.  
XX PR 07-JUL-2000; 2000US-0216747P.  
XX XX (CORI-) CORIXA CORP.  
XX PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
XX WPI; 2001-616774/71.  
XX DR N-PSDB; AAS59513.  
XX PT Propionibacterium acnes polypeptides and nucleic acids useful for  
XX PT vaccinating against and diagnosing infections, especially useful for  
XX PT treating acne vulgaris.

XX Example 1; SEQ ID NO 2127; 1069pp; English.  
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
XX polypeptides. The proteins and their associated DNA sequences are used in  
XX the treatment, prevention and diagnosis of medical conditions caused by  
XX P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
XX pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
XX P. acnes is also involved in infections of bone, joints and the central  
XX nervous system, however it is particularly involved in the inflammatory  
XX lesions associated with acne vulgaris. A method for detecting the  
XX presence or absence of P. acnes in a patient comprises contacting a  
XX sample with a binding agent that binds to the proteins of the invention  
XX and determining the amount of bound protein in the sample. The  
XX polypeptides may be used as antigens in the production of antibodies  
XX specific for P. acnes proteins. These antibodies can be used to  
XX downregulate expression and activity of P. acnes polypeptides and  
XX therefore treat P. acnes infections. The antibodies may also be used as  
XX diagnostic agents for determining P. acnes presence, for example, by  
XX enzyme linked immunosorbent assay (ELISA). Note: the sequence data for  
XX this patent did not form part of the printed specification, but was  
XX obtained in electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences

Sequence 535 AA;

Query Match 71.4%; Score 35; DB 4; Length 535;  
Best Local Similarity 70.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAVNHQKSA 10  
Db 488 AAVNHQKSA 497

RESULT 10  
ABM37451  
ID ABM37451 standard; protein; 535 AA.  
XX AC ABM37451;  
XX DT 20-OCT-2003 (first entry)  
XX DE Propionibacterium acnes predicted ORF-encoded polypeptide #2127.  
XX KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;  
XX KW immunostimulant; immune response; vaccine.  
XX OS Propionibacterium acnes.

XX WO2003033515-A1.  
XX PN 24-APR-2003.  
XX PF 11-OCT-2002; 2002WO-US032727.  
XX PR 15-OCT-2001; 2001US-00978825.  
XX XX (CORI-) CORIXA CORP.  
XX PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;  
XX PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;  
XX PI Barth B, Vallieue-Douglas J;  
XX WPI; 2003-381789/36.  
XX DR N-PSDB; ACF64442.  
XX PT New Propionibacterium acnes polypeptides and polynucleotides encoding the  
XX PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,  
XX PT or for stimulating an immune response specific for a P. acnes protein.  
XX Example 1; SEQ ID NO 2127; 1481pp; English.

XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)  
XX encoding a Propionibacterium acnes protein. The invention also relates to  
XX polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to  
XX immunogenic fragments of P. acnes polypeptides. The invention  
XX additionally encompasses expression vectors and host cells comprising a  
XX polynucleotide of the invention; antibodies against polypeptides of the  
XX invention; fusion proteins comprising a polypeptide of the invention; a  
XX method for stimulating an immune response specific for a P. acnes  
XX polypeptide and an isolated T cell population comprising T cells prepared  
XX via this method; a vaccine composition comprising P. acnes polypeptides,  
XX polynucleotides, antibodies, fusion proteins, T cell populations, or  
XX antigen-presenting cells that express the polypeptide; a method and kit  
XX for detecting or determining the presence or absence of P. acnes in a  
XX patient; and a method for inhibiting the development of P. acnes in a  
XX patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion  
XX proteins, T cell populations or antigen-presenting cells that express the  
XX polypeptides are useful for diagnosing, preventing or treating acne  
XX vulgaris, or for stimulating an immune response specific for a P. acnes  
XX protein. The polynucleotides can also be used as probes or primers for  
XX nucleic acid hybridisation. The vaccine composition is useful for the  
XX stimulation of an immune response against P. acnes, or for treating acne,  
XX and the kit is useful for performing a diagnostic assay. The present  
XX sequence represents a polypeptide predicted to be encoded by an ORF (open  
XX reading frame) contained within the P. acnes polynucleotides of the  
XX invention. Note: The sequence data for this patent did not form part of

CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 535 AA;

Query Match 71.4%; Score 35; DB 6; Length 535;  
 Best Local Similarity 70.0%; Pred. No. 2.2e+02;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAVNHQKSA 10  
 ||||| :  
 Db 488 AAVNHSRTAA 497

RESULT 11  
 ADD31051  
 ID ADD31051 standard; protein; 664 AA.

XX  
 AC ADD31051;

XX DT 15-JAN-2004 (first entry)

XX Plant yield-related protein from clone G1079.

XX DE transcription factor; transgenic plant; growth rate; senescence;  
 XX KW seed germination rate; plant vigor; seedling vigor.

XX OS Arabidopsis thaliana.

XX PN WO2003013227-A2.

XX PD 20-FEB-2003.

XX PF 09-AUG-2002; 2002WO-US025805.

XX PR 09-AUG-2001; 2001US-0310847P.

XX PR 19-NOV-2001; 2001US-0336049P.

XX PR 11-DEC-2001; 2001US-0338692P.

XX PR 14-JUN-2002; 2002US-00171468.

XX (MEND-) MENDEL BIOTECHNOLOGY INC.

XX Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JE;

XX Pilgrim ML, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;

XX Broun PE;

XX WPI: 2003-248221/24.

XX N-PSDB; ADD31050.

XX New plant transcription factor polynucleotides and polypeptides, useful

XX in producing transgenic plants with commercially valuable properties,

XX such as an alteration in a plant growth characteristic, e.g. growth rate

XX or apomixis.

XX Disclosure; SEQ ID NO 1080; 454pp; English.

XX The invention relates to a number of isolated Arabidopsis thaliana cDNA

XX sequences and their encoded proteins which are especially transcription

XX factor related cDNA's and proteins. The isolated or recombinant plant

XX transcription factor polynucleotides and polypeptides are useful in

XX producing transgenic plants with commercially valuable properties, i.e.

XX modified or altered desirable traits as compared to a reference plant,

XX such as an alteration in a plant growth characteristic, e.g. growth rate,

XX germination rate of seeds, vigor of plants and seedlings, or leaf and

XX flower senescence. Sequence information related to the polynucleotides

XX and polypeptides can also be used in bioinformatic search methods. The

XX transgenic plant is useful for growing a progeny plant from a parent

XX plant. This sequence represents one of the proteins of the invention.

XX Sequence 664 AA;

XX Query Match 71.4%; Score 35; DB 7; Length 664;

XX Best Local Similarity 70.0%; Pred. No. 2.8e+02;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAVNHQKSA 10  
 ||||| :  
 Db 359 AKVEHERKSA 368

RESULT 12

ADI44067

ID ADI44067 standard; protein; 664 AA.

XX AC ADI44067;

XX DT 22-APR-2004 (first entry)

XX DE Plant transcription factor related polypeptide #1622.

XX KW transgenic; plant; enhanced tolerance to abiotic stress;

XX KW phosphate tolerance; hormone sensitivity; disease resistance;

XX KW sugar sensing; flowering; flower structure; stem bifurcation;

XX KW branching pattern; apical dominance; trichome; stem morphology;

XX KW root growth; root hair; seed development; cell proliferation;

XX KW cell differentiation; premature senescence; necrosis; plant size;

XX KW leaf morphology; seed morphology; seed biochemistry; root anthocyanin;

XX KW plant anthocyanin; light response; shade avoidance; bioinformatic;

XX KW transcription factor; ds.

XX OS Unidentified.

XX PN US2004019927-A1.

XX PD 29-JAN-2004.

XX PF 25-FEB-2003; 2003US-00374780.

XX PR 18-APR-2001; 2001US-00837944.

XX (SHER/) SHERMAN B K.

XX (RIEC/) RIECHMANN J L.

XX (JIANG/) JIANG C.

XX (HEAR/) HEARD J E.

XX (HAAK/) HAAKE V.

XX (CREE/) CREELMAN R A.

XX (RATC/) RATCLIFFE O.

XX (ADAM/) ADAM L J.

XX (REUB/) REUBER T L.

XX (KEDD/) KEDDIE J.

XX (BROU/) BROUN P E.

XX (PILG/) PILGRIM M L.

XX (DUBE/) DUBELL A N.

XX (PINE/) PINEDA O.

XX (YUGG/) YU G.

XX Sherman BK, Riechmann JL, Jiang C, Heard JE, Haaake V;

XX Creelman RA, Ratcliffe O, Adam LJ, Reuber TL, Keddle J, Broun PE;

XX Pilgrim ML, Dubell AN, Pineda O, Yu G;

XX WPI: 2004-132245/13.

XX N-PSDB; ADI44066.

XX New transgenic plant comprising a recombinant polynucleotide of any one  
 of more than 500 nucleotide sequences, useful in bioinformatic search  
 methods.

XX Disclosure; SEQ ID NO 2530; 435pp; English.

XX The invention describes a transgenic plant comprising a recombinant  
 polynucleotide of any one of more than 500 nucleotide sequences fully  
 defined in the specification or its complement. The method of the  
 invention can be used to produce a plant having altered traits such as:  
 enhanced tolerance to abiotic stress; phosphate tolerance; hormone  
 sensitivity; disease resistance; sugar sensing; early or late flowering;  
 altered flower structure, change in stem bifurcations, altered branching

CC pattern, reduced apical dominance, reduced trichome density; lack of  
 CC trichomes; reduced ectopic trichome development; altered trichome  
 CC development; increase in trichome number; altered stem morphology;  
 CC increased root growth; increased root hairs; altered seed development;  
 CC altered cell proliferation or cell differentiation; rapid development;  
 CC premature senescence; increased necrosis; increase in seedling or plant  
 CC size; decreased plant size; leaf morphology; seed morphology; seed  
 CC biochemistry; increase in root anthocyanins; increase in plant  
 CC anthocyanins; or alteration in light response or shade avoidance. The  
 CC transgenic plant, polynucleotides and polypeptides are useful in  
 CC bioinformatic search methods. This sequence represents a plant  
 CC transcription factor related polynucleotide.

XX SQ Sequence 664 AA;

Query Match 71.4%; Score 35; DB 8; Length 664;  
 Best Local Similarity 70.0%; Pred. No. 2.8e+02;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAVNHQKSA 10  
 | | | | |  
 Db 359 AKVEHERKSA 368

RESULT 13

ADN96832  
 ID ADN96832 standard; protein; 5509 AA.

XX AC ADN96832;  
 XX DT 01-JUL-2004 (first entry)  
 XX DE Bugula bryostatin gene cluster ORF1 protein SeqID 43.  
 XX KW polypeptide; bryopyran ring; bryostatin; polyketide synthase; PKS; toxic;  
 XX KW cytostatic; immunomodulatory; protein therapy; cancer; metastasis.  
 XX OS Bugula.  
 XX PN WO2003099219-A2.  
 XX PD 04-DEC-2003.  
 XX PF 20-MAY-2003; 2003WO-US016299.  
 XX PR 20-MAY-2002; 2002US-0382181P.  
 XX PA (REGC ) UNIV CALIFORNIA.  
 XX PA (HAYG/) HAYGOOD M.  
 XX PA (HILD/) HILDEBRAND M.  
 XX PA (ANDE/) ANDERSON C.  
 XX PA (WAGG/) WAGGONER L E.  
 XX PA (SHER/) SHERMAN D H.  
 XX PA (LIUH/) LIU H.  
 XX PI Haygood M, Hildebrand M, Anderson C, Waggoner LE, Sherman DH;  
 XX PI Liu H;  
 XX DR WPI; 2004-053143/05.  
 XX PT New compositions comprising a polyketide synthase or nucleic acid  
 XX PT encoding the polyketide synthase, useful in biosynthesizing polyketides,  
 XX PT bryopyran rings and bryostatins having anti-cancer or antimetastatic  
 XX PT activity.  
 XX PS Disclosure; Fig 25c; 342pp; English.  
 XX CC This invention relates to a novel composition that comprises at least one  
 CC polypeptide that catalyses the one step synthesis of a polyketide or  
 CC bryopyran ring. Specifically, it refers to nucleic acid molecules derived  
 CC from marine organisms that encode enzymes that catalyse the synthesis of  
 CC bioactive compounds such as polyketides and bryostatins that are based on  
 CC the bryopyran ring structure. The present invention describes methods for

CC the use of polyketide synthases (PKSs) to generate toxic polyketides that  
 CC exhibit cytostatic and immunomodulatory activities, such that they can be  
 CC used for protein therapy in the treatment of cancer and metastasis. This  
 CC polypeptide is a Bugula bryostatin biosynthetic gene cluster encoded  
 CC protein of the invention. NOTE: This SeqID number is also attributed to a  
 XX primer given in figure 28b.

XX SQ Sequence 5509 AA;

Query Match 71.4%; Score 35; DB 8; Length 5509;  
 Best Local Similarity 66.7%; Pred. No. 2.6e+03;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAVNHQKKS 9  
 : | | | | | :  
 Db 3131 SAVNHQKKN 3139

RESULT 14

ADN96829  
 ID ADN96829 standard; protein; 6940 AA.

XX AC ADN96829;  
 XX DT 01-JUL-2004 (first entry)  
 XX DE Bugula bryostatin gene cluster encoded protein (reading frame 2).  
 XX KW polypeptide; bryopyran ring; bryostatin; polyketide synthase; PKS; toxic;  
 XX KW cytostatic; immunomodulatory; protein therapy; cancer; metastasis.  
 XX OS Bugula.  
 XX FH Key Location/Qualifiers  
 FT Misc-difference 1..6940  
 FT /label= Xaa  
 FT /note= "Xaa = in frame stop codons"  
 XX PN WO2003099219-A2.  
 XX PD 04-DEC-2003.  
 XX PF 20-MAY-2003; 2003WO-US016299.  
 XX PR 20-MAY-2002; 2002US-0382181P.  
 XX PA (REGC ) UNIV CALIFORNIA.  
 XX PA (HAYG/) HAYGOOD M.  
 XX PA (HILD/) HILDEBRAND M.  
 XX PA (ANDE/) ANDERSON C.  
 XX PA (WAGG/) WAGGONER L E.  
 XX PA (SHER/) SHERMAN D H.  
 XX PA (LIUH/) LIU H.  
 XX PI Haygood M, Hildebrand M, Anderson C, Waggoner LE, Sherman DH;  
 XX PI Liu H;  
 XX DR WPI; 2004-053143/05.  
 XX DR N-PSDB; ADN96827.  
 XX PT New compositions comprising a polyketide synthase or nucleic acid  
 XX PT encoding the polyketide synthase, useful in biosynthesizing polyketides,  
 XX PT bryopyran rings and bryostatins having anti-cancer or antimetastatic  
 XX PT activity.  
 XX PS Disclosure; Fig 25a; 342pp; English.  
 XX CC This invention relates to a novel composition that comprises at least one  
 CC polypeptide that catalyses the one step synthesis of a polyketide or  
 CC bryopyran ring. Specifically, it refers to nucleic acid molecules derived  
 CC from marine organisms that encode enzymes that catalyse the synthesis of  
 CC bioactive compounds such as polyketides and bryostatins that are based on  
 CC the bryopyran ring structure. The present invention describes methods for

CC the use of polyketide synthases (PKSs) to generate toxic polyketides that  
 CC exhibit cytostatic and immunomodulatory activities, such that they can be  
 CC used for protein therapy in the treatment of cancer and metastasis. This  
 CC polypeptide is a Bugula bryostatins biosynthetic gene cluster encoded  
 CC protein (reading frame 2) of the invention.  
 XX

SQ Sequence 6940 AA;

Query Match 71.4%; Score 35; DB 8; Length 6940;

Best Local Similarity 66.7%; Pred. No. 3.4e+03;

Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAVNHQRKS 9

:|||||:

Db 3321 SAVNHQQRN 3329

RESULT 15

AD530840  
 ID ADS30840 standard; protein; 284 AA.

XX AC

XX ADS30840;

XX DT 02-DEC-2004 (first entry)

XX DE Bacterial polypeptide #19873.

XX Recombinant DNA construct; transformed plant; improved plant property;  
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
 KW pathogen tolerance; pest tolerance; plant disease resistance;  
 KW cell cycle pathway modification; plant growth regulator;  
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
 KW bacterial polypeptide.

XX OS Bacteria.

XX US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.

XX (HINK/) HINKLE G J.

XX (SLAT/) SLATER S C.

XX (CHEN/) CHEN X.

XX (GOLD/) GOLDMAN B S.

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide  
 PT for expression of a polynucleotide encoding a polypeptide from a  
 PT microbial source, useful for producing plants with improved properties.

PS Claim 1; SEQ ID NO 19873; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a  
 CC promoter functional in a plant cell, where the promoter is positioned to  
 CC provide for expression of a polynucleotide encoding a polypeptide from a  
 CC microbial source. The invention also relates to a transformed plant  
 CC comprising the recombinant DNA construct and a method of producing a  
 CC transformed plant having an improved property. The plant is a crop plant  
 CC such as maize or soybean. The method of producing a transformed plant  
 CC having an improved property comprises transforming a plant with the  
 CC recombinant DNA construct and growing the transformed plant, where the  
 CC polynucleotide or polypeptide is useful for improving plant properties.  
 CC The recombinant DNA construct is useful for producing plants with  
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,

CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
 CC increased resistance to plant disease, better growth rate by modification  
 CC of the cell cycle pathway with plant growth regulators, increased rate of  
 CC homologous recombination, modified seed oil or protein yield and/or  
 CC content, improved yield by modification of carbohydrate, nitrogen or  
 CC phosphorus use and/or uptake, by modification of photosynthesis or by  
 CC providing improved plant growth and development under at least one stress  
 CC condition, improved lignin production or improved galactomannan  
 CC production. This sequence represents a bacterial polypeptide used in the  
 CC scope of the invention. Note: The sequence data for this patent did not  
 CC form part of the printed specification but was obtained in electronic  
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 284 AA;

Query Match 69.4%; Score 34; DB 8; Length 284;

Best Local Similarity 81.8%; Pred. No. 1.8e+02;

Matches 9; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

Qy 1 AAVNH--QRKS 9

Db 126 AAVNHFKQRKS 136

Search completed: May 3, 2006, 19:01:23

Job time : 8.24897 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2006, 19:01:48 ; Search time 0.751029 Seconds  
(without alignments)  
1281.133 Million cell updates/sec

Title: US-10-525-567-1  
Perfect score: 49  
Sequence: 1 AAVNHQKSA 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_80.\*  
1: Pirl.\*  
2: Pirl.\*  
3: Pirl.\*  
4: Pirl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	100.0	467	2 S39193	royal jelly protei
2	38	77.6	169	2 AF2630	hypothetical prote
3	37	75.5	194	2 B86073	hypothetical prote
4	37	75.5	194	2 D12226	Molybdopterin-guan
5	37	75.5	194	2 S40803	molybdopterin-guan
6	37	75.5	532	2 T49824	hypothetical prote
7	35	71.4	664	2 T01368	hypothetical prote
8	34	69.4	253	2 A88955	protein K04F1.1 [l
9	34	69.4	400	2 H69009	hypothetical prote
10	34	69.4	448	2 G87664	glycine cleavage s
11	34	69.4	479	2 S85466	xylose isomerase (
12	33	67.3	211	2 T47915	LIM domain protein
13	33	67.3	418	2 G87581	conjugal transfer
14	33	67.3	810	2 A81965	hemeoglobin-hapto
15	32	65.3	236	2 T45067	hypothetical prote
16	32	65.3	296	2 E81132	transcription regu
17	32	65.3	296	2 S24954	regulatory protein
18	32	65.3	305	2 A80863	positive regulator
19	32	65.3	305	2 F85932	glycine cleavage s
20	32	65.3	305	2 I41065	positive regulator
21	32	65.3	305	2 D91087	glycine cleavage s
22	32	65.3	305	2 AD0126	hypothetical prote
23	32	65.3	321	2 F82652	glycine betaine-bi
24	32	65.3	331	2 AC0842	adenosylhomocyste
25	32	65.3	527	1 S01302	phosphoprotein pho
26	32	65.3	692	1 PABY12	forked protein 5.6
27	32	65.3	1449	2 S57237	hypothetical prote
28	31	63.3	168	2 T17204	conserved hypothet
29	31	63.3	185	1 B69374	

30	31	63.3	275	2 A11965	hypothetical prote
31	31	63.3	433	2 T06407	monodehydroascorba
32	31	63.3	434	2 JU0182	monodehydroascorba
33	31	63.3	434	2 T47545	monodehydroascorba
34	31	63.3	470	2 S71355	glucocorticoid-att
35	31	63.3	472	2 A55508	interferon alpha I
36	31	63.3	524	2 T20872	hypothetical prote
37	31	63.3	626	2 I49100	macd6 precursor -
38	31	63.3	681	2 E64598	hypothetical prote
39	31	63.3	693	2 T19551	mucin-like protein
40	31	63.3	811	2 T39336	DEC1 protein homol
41	31	63.3	1024	2 AB0347	AcroB/AcroP/AcroF fam
42	31	63.3	1047	2 T46489	hypothetical prote
43	31	63.3	1300	2 I53799	CGI protein - huma
44	31	63.3	1350	2 S00647	finger protein - A
45	31	63.3	1356	2 S32763	kinectin 1 - human

ALIGNMENTS

RESULT 1

S39193  
royal jelly protein RJP57-1 - honeybee  
C;Species: Apis mellifera (honeybee)  
C;Date: 05-Mar-1995 #sequence\_revision 12-May-1995 #text\_change 12-May-1995  
C;Accession: S39193  
R;Klaudiny, J.; Hanes, J.; Kulifajova, J.; Albert, S.; Simuth, J.  
submitted to the EMBL Data Library, September 1993  
A;Description: Molecular cloning of two cDNAs from the head of the nurse honey bee (Apis mellifera) cDNA library.  
A;Reference number: S39193  
A;Accession: S39193  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-467 <KLA>  
A;Cross-references: UNIPARC:UPI0000179A48; EMBL:Z26318

Query Match 100.0%; Score 49; DB 2; Length 467;  
Best Local Similarity 100.0%; Pred. No. 0.017;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAVNHQKSA 10  
| | | | | | | | | |  
Db 21 AAVNHQKSA 30

RESULT 2

AF2630  
hypothetical protein Atu0441 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)  
C;Species: Agrobacterium tumefaciens  
C;Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
C;Accession: AF2630  
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.; Wood, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan, P.; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, A.; Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A;Reference number: AB2577; MUID:21608550; PMID:11743193  
A;Accession: AF2630  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-169 <KUR>  
A;Cross-references: UNIPROT:Q8UI59; UNIPARC:UPI000000D17ED; GB:AE008688; PIDN:AAL41460.1;  
A;Experimental source: strain C58 (Dupont)  
C;Genetics:  
A;Gene: Atu0441  
A;Map position: circular chromosome

Query Match 77.6%; Score 38; DB 2; Length 169;  
Best Local Similarity 70.0%; Pred. No. 1.3;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAVNHQKSA 10  
:|||||:|  
Db 17 SAVNHSRSA 26

## RESULT 3

hypothetical protein mobA [imported] - Escherichia coli (strain O157:H7, substrain EDL933)  
C:Species: Escherichia coli  
C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 05-Oct-2004  
C:Accession: B86073  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoudis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: B86073  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-194 <STO>  
A:Cross-references: UNIPARC:UPI00001659A4; GB:AE005174; NID:g12518733; PIDN:AAG59046.1;  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: mobA  
C:Superfamily: molybdopterin guanylyltransferase MobA

Query Match 75.5%; Score 37; DB 2; Length 194;  
Best Local Similarity 70.0%; Pred. No. 2.5;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAVNHQKSA 10  
:|||||:|  
Db 111 ARLNHQKDA 120

## RESULT 4

D91226  
Molybdopterin-guanine dinucleotide biosynthesis protein A [similarity] - Escherichia coli  
C:Species: Escherichia coli  
C>Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 05-Oct-2004  
C:Accession: D91226  
R:Hayaishi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: D91226  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-194 <HAY>  
A:Cross-references: UNIPROT:P58221; UNIPARC:UPI000012F34D; GB:BA000007; PIDN:BA038203.1;  
A:Experimental source: strain O157:H7, substrain RIMD 0509952  
C:Genetics:  
A:Gene: EceA780  
C:Superfamily: molybdopterin guanylyltransferase MobA

Query Match 75.5%; Score 37; DB 2; Length 194;  
Best Local Similarity 70.0%; Pred. No. 2.5;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAVNHQKSA 10  
:|||||:|  
Db 111 ARLNHQKDA 120

## RESULT 5

S40803  
molybdopterin-guanine dinucleotide biosynthesis protein A - Escherichia coli (strain K-12)  
N:Alternate names: mob protein  
C:Species: Escherichia coli  
C>Date: 06-Oct-1994 #sequence\_revision 26-May-1995 #text\_change 05-Oct-2004  
C:Accession: S40803; E65190

R;Plunkett III, G.; Burland, V.; Daniels, D.L.; Blattner, F.R.  
Nucleic Acids Res. 21, 3391-3398, 1993  
A>Title: Analysis of the Escherichia coli genome. III. DNA sequence of the region from 8;  
A:Reference number: S40802; MUID:93347969; PMID:8346018

A:Accession: S40803  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA

A:Residues: 1-194 <PLU>  
A:Cross-references: UNIPROT:P32173; UNIPARC:UPI0000111770; EMBL:L19201; NID:g304961; PIDN:  
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1993  
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col  
A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997

A>Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: E65190

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-194 <BLAT>

A:Cross-references: UNIPARC:UPI0000111770; GB:AE000461; GB:U00096; NID:g2367318; PIDN:AAK

A:Experimental source: strain K-12, substrain MG1655

C:Genetics:

A:Gene: mobA

A:Start codon: GTG

C:Superfamily: molybdopterin guanylyltransferase MobA

Query Match 75.5%; Score 37; DB 2; Length 194;  
Best Local Similarity 70.0%; Pred. No. 2.5;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAVNHQKSA 10  
:|||||:|  
Db 111 ARLNHQKDA 120

## RESULT 6

T49824  
hypothetical protein B24H17.100 [imported] - Neurospora crassa  
C:Species: Neurospora crassa  
C>Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000  
C:Accession: T49824  
R;Schulte, U.; Align, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,  
submitted to the Protein Sequence Database, May 2000  
A:Reference number: Z25022  
A:Accession: T49824  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-532 <SCH>  
A:Cross-references: UNIPARC:UPI0000179E42; EMBL:AL356815; GSPDB:GN00116; NCSP:B24H17.100  
A:Experimental source: BAC clone B24H17; strain OR74A  
C:Genetics:  
A:Gene: NCSP:B24H17.100  
A:Map position: 6  
A:Introns: 14/1; 169/1; 284/3; 429/1; 496/3

Query Match 75.5%; Score 37; DB 2; Length 532;  
Best Local Similarity 100.0%; Pred. No. 7.2;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VNHQKS 9  
:|||||:|  
Db 221 VNHQKS 227

## RESULT 7

T01368  
hypothetical protein At2g34670 [imported] - Arabidopsis thaliana  
N:Alternate names: hypothetical protein T29F13.12  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004  
C:Accession: T01368; E84759  
R;Roundley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul,  
submitted to the EMBL Data Library, May 1998

A;Description: Arabidopsis thaliana chromosome II BAC T29F13 genomic sequence.  
A;Reference number: Z14179  
A;Accession: T01368  
A;Status: translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-664 <ROU>  
A;Cross-references: UNIPROT:O64589; UNIPARC:UPI00000A0612; EMBL:AC003096; NID:g3132469;  
A;Experimental source: cultivar Columbia  
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402: 761-768, 1999  
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A;Reference number: A84420; MUID:20083487; PMID:10617197  
A;Accession: E84759  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-664 <STO>  
A;Cross-references: UNIPARC:UPI00000A0612; GB:AE002093; NID:g3132478; PIDN:AAC16267.1; G  
C;Genetics:  
A;Gene: At2g34670; T29F13.12  
A;Map position: 2  
A;Introns: 294/1; 354/3; 429/2; 532/3

Query Match 71.4%; Score 35; DB 2; Length 664;  
Best Local Similarity 70.0%; Pred. No. 24;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAVNHQRKSA 10  
| | | | |  
Db 359 AKVEHRKSA 368

RESULT 8  
A88955  
Protein K04F1.1 [imported] - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C;Accession: A88955  
R;Anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog  
A;Reference number: A75000; MUID:99069613; PMID:9851916  
A;Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/projects/C\_ele  
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A;Accession: A88955  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-253 <STO>  
A;Cross-references: UNIPROT:Q9TXL6; UNIPARC:UPI000007C0A4; GB:chr\_V; PIDN:AAC78175.1; PI  
C;Genetics:  
A;Gene: K04F1.1  
A;Map position: 5

Query Match 69.4%; Score 34; DB 2; Length 253;  
Best Local Similarity 60.0%; Pred. No. 14;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAVNHQRKSA 10  
| | | | |  
Db 3 AQINHYRKSS 12

RESULT 9  
H69009  
Hypothetical protein MTH1076 - Methanobacterium thermoautotrophicum (strain Delta H)  
C;Species: Methanobacterium thermoautotrophicum  
C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C;Accession: H69009  
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;  
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwnani, N.  
Kl, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997

A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct  
A;Reference number: A69000; MUID:98037514; PMID:9371463  
A;Accession: H69009  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-400 <MTH>  
A;Cross-references: UNIPROT:O27148; UNIPARC:UPI0000062B8B; GB:AE000879; GB:AE000666; NID  
A;Experimental source: strain Delta H  
C;Genetics:  
A;Gene: MTH1076

Query Match 69.4%; Score 34; DB 2; Length 400;  
Best Local Similarity 75.0%; Pred. No. 23;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VNHQRKSA 10  
| | | | |  
Db 182 VNHQRKSA 189

RESULT 10  
G87664  
glycine cleavage system P protein, subunit 1 [imported] - Caulobacter crescentus  
C;Species: Caulobacter crescentus  
C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
C;Accession: G87664  
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A;Title: Complete Genome Sequence of Caulobacter crescentus.  
A;Reference number: A87249; MUID:21173698; PMID:11259647  
A;Accession: G87664  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-448 <STO>  
A;Cross-references: UNIPROT:Q9A353; UNIPARC:UPI000000C7A48; GB:AE005673; NID:g13425055; P  
C;Genetics:  
A;Gene: CC3353  
C;Superfamily: Aquifex aeolicus glycine dehydrogenase (decarboxylating)

Query Match 69.4%; Score 34; DB 2; Length 448;  
Best Local Similarity 70.0%; Pred. No. 26;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAVNHQRKSA 10  
| | | | |  
Db 352 AAVNHQRKSA 361

RESULT 11  
S65466  
xylose isomerase (SC 5.3.1.5) - barley  
C;Species: Hordeum vulgare (barley)  
C;Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
C;Accession: S65466; S65467; S65405  
R;Kristo, P.; Saarelainen, R.; Fagerstrom, R.; Aho, S.; Korhola, M.  
Eur. J. Biochem. 237, 240-246, 1996  
A;Title: Protein purification, and cloning and characterization of the cDNA and gene for  
A;Reference number: S65405; MUID:96203931; PMID:8620879  
A;Accession: S65466  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-479 <KRI>  
A;Cross-references: UNIPROT:Q40082; UNIPARC:UPI000016DDE9; EMBL:X95256; NID:g1396806; PI  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1996  
A;Accession: S65467  
A;Molecule type: mRNA  
A;Residues: 1-442; AGKGDETELEKKALENGEPTVPSGQ, 472-479 <KRF>  
A;Cross-references: UNIPARC:UPI00000176187; EMBL:X95257  
A;Accession: S65405  
A;Molecule type: protein  
A;Residues: 54-62; 69-86; 97-108; 128-135; 193-204; 235-246; 300-307; 321-330; 371-379 <KRW>

A;Cross-references: UNIPARC:UPI0000176188; UNIPARC:UPI0000176189; UNIPARC:UPI000017618A;  
18F; UNIPARC:UPI0000176190  
C;Genetics:  
A;Introns: 20/3; 57/3; 82/3; 126/1; 156/3; 173/3; 200/2; 218/3; 252/3; 270/1; 285/2; 304  
C;Superfamily: xylose isomerase  
C;Keywords: intramolecular oxidoreductase; isomerase

Query Match 59.4%; Score 34; DB 2; Length 479;  
Best Local Similarity 75.0%; Pred. No. 28;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAVNHQK 8  
|||:|:  
Db 258 AAVNHKKK 265

RESULT 12  
T47915  
LIM domain protein - Arabidopsis thaliana  
N;Alternate names: protein T20K12.130  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 05-Oct-2004  
C;Accession: T47915  
R;De Haan, M.; Maarse, A.C.; Grivell, L.A.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quest  
submitted to the Protein Sequence Database, January 2000  
A;Reference number: Z24480  
A;Accession: T47915  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-211 <DEH>  
A;Cross-references: UNIPROT:Q9W2E1; UNIPARC:UPI000009EDBC; EMBL:AL137898  
A;Experimental source: cultivar Columbia; BAC clone T20K12  
C;Genetics:  
A;Map position: 3  
A;Introns: 46/3; 80/1; 89/3  
A;Note: T20K12.130  
C;Superfamily: cysteine and glycine-rich protein; LIM metal-binding repeat homology

Query Match 67.3%; Score 33; DB 2; Length 211;  
Best Local Similarity 60.0%; Pred. No. 19;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAVNHQKSA 10  
|||:|:  
Db 173 AAVNHRRSTA 182

RESULT 13  
G87581  
conjugal transfer protein TrbI [imported] - Caulobacter crescentus  
C;Species: Caulobacter crescentus  
C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
C;Accession: G87581  
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.;  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A;Title: Complete Genome Sequence of Caulobacter crescentus.  
A;Reference number: A87249; MUID:21173698; PMID:11259647  
A;Accession: G87581  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-418 <STO>  
A;Cross-references: UNIPROT:Q9A4Y7; UNIPARC:UPI000000C77E7; GB:AE005673; NID:G13424269; F  
C;Genetics:  
A;Gene: CC2685

Query Match 67.3%; Score 33; DB 2; Length 418;  
Best Local Similarity 62.5%; Pred. No. 40;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAVNHQK 8  
|||:|:  
Db 173 AAVNHRRSTA 182

Db 114 AVINHQR 121

RESULT 14  
A81965  
hemeoglobin-haptoglobin-utilization protein NMA0474 [imported] - Neisseria meningitidis  
C;Species: Neisseria meningitidis  
C;Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 09-Jul-2004  
C;Accession: A81965  
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morell  
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,  
Nature 404, 502-506, 2000  
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.  
A;Reference number: A81775; MUID:20222556; PMID:10761919  
A;Accession: A81965  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-810 <PAR>  
A;Cross-references: UNIPROT:Q5JWA2; UNIPARC:UPI000012CB78; GB:AL162753; GB:AL157959; NID  
A;Experimental source: serogroup A, strain Z2491  
C;Genetics:  
A;Gene: hpuB; NMA0474

Query Match 67.3%; Score 33; DB 2; Length 810;  
Best Local Similarity 77.8%; Pred. No. 80;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 AVNHQKSA 10  
|||:|:  
Db 160 AVNYQKSA 168

RESULT 15  
T45067  
hypothetical protein [imported] - Helicobacter pylori plasmid PHPM180  
C;Species: Helicobacter pylori  
C;Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004  
C;Accession: T45067  
R;Minnis, J.; Knesek, J.; Taylor, T.; McIntire, S.A.  
submitted to the EMBL Data Library, July 1994  
A;Description: Cloning and DNA sequence analysis of a plasmid from Helicobacter pylori.  
A;Reference number: Z22904  
A;Accession: T45067  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-236 <MIN>  
A;Cross-references: UNIPROT:Q48249; UNIPARC:UPI000000B75FF; EMBL:U12689; PIDN:AAA20889.1  
A;Experimental source: strain HPM180  
C;Genetics:  
A;Genome: plasmid PHPM180

Query Match 65.3%; Score 32; DB 2; Length 236;  
Best Local Similarity 62.5%; Pred. No. 36;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 VNHQKSA 10  
|||:|:  
Db 10 LNHERKNA 17

Search completed: May 3, 2006, 19:10:17  
Job time : 2.75103 secs



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2006, 18:54:38 ; Search time 4.62963 Seconds

(without alignments)  
1523.941 Million cell updates/sec

Title: US-10-525-567-1

Perfect score: 49

Sequence: 1 AAVNHORKSA 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 05.80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	100.0	31	Q86R92	apis mellif
2	49	100.0	544	MRJP3	apis mellif
3	40	81.6	344	Q7P382	fusobacteri
4	38	77.6	169	Q8UI59	agrobacteri
5	37	75.5	194	MOBA_ECO57	P58221 escherichia
6	37	75.5	194	MOBA_ECOLI	P32173 escherichia
7	37	75.5	194	Q83PF7	shigella fl
8	36	73.5	135	Q8BN36	mus musculus
9	36	73.5	347	Q8BN36_MOUSE	Q97459 drosophila
10	36	73.5	347	Q9VXN1	drosophila
11	36	73.5	375	Q8SX00	drosophila
12	36	73.5	503	Q9V7R7	drosophila
13	35	71.4	277	Q7NAK1	mycoplasma
14	35	71.4	304	Q4HFA2	campylobact
15	35	71.4	357	Q6MIG1	bdellovibri
16	35	71.4	369	Q6AJ53	desulfotale
17	35	71.4	462	Q6ABV2	propionibac
18	35	71.4	561	Q8GW98	arabidopsis
19	35	71.4	650	Q892X9	clostridium
20	35	71.4	664	Q64589	arabidopsis
21	35	71.4	689	Q8X179	emeridella
22	35	71.4	724	Q6FU10	candida gla
23	35	71.4	922	Q887R0	pseudomonas
24	35	71.4	1307	Q7SMH0	neurospora
25	34	69.4	214	Q8FJ67	escherichia
26	34	69.4	244	Q9TXL6	caenorhabdi
27	34	69.4	284	Q5BQ55	aspergillus
28	34	69.4	306	Q7N8R6	photorhabd
29	34	69.4	354	Q5Q3W2	macaca neme
30	34	69.4	374	Q822Z0	chlamydomo
31	34	69.4	387	Q9H5E5	homo sapien

32	34	69.4	398	2	Q76KC7	ORYLA
33	34	69.4	398	2	Q90W00	ORYLA
34	34	69.4	400	2	O27148	METH
35	34	69.4	410	2	O821G0	CHLCV
36	34	69.4	412	2	Q41P63	GIBBERE
37	34	69.4	419	2	Q53FW3	HUMAN
38	34	69.4	431	2	O41E78	GIBBE
39	34	69.4	448	1	GCSPA	CAUCR
40	34	69.4	452	1	XYLW	HORVU
41	34	69.4	479	1	XYLA	HORVU
42	34	69.4	556	2	Q6CZ20	ERWCT
43	34	69.4	562	2	Q6ZN84	HUMAN
44	34	69.4	563	2	Q4NAP4	9MICC
45	34	69.4	1208	2	Q4PELO	USTWA

Q76kc7 oryzias lat  
Q90w00 oryzias lat  
O27148 methanobact  
O821g0 chlamydomo  
O41p63 gibberella  
Q53fw3 homo sapien  
O41e78 gibberella  
Q9a353 caulobacter  
Q720w0 trypanosoma  
Q40082 hordeum vul  
O6cz20 erwinia car  
Q6zn84 homo sapien  
Q4na44 arthrobacte  
Q4pe10 ustilago ma

#### ALIGNMENTS

##### RESULT 1

Q86R92 APIME  
ID Q86R92 APIME PRELIMINARY; PRT; 31 AA.  
AC Q86R92;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Major royal jelly protein MRJP3 (Fragment).  
GN Name=MRJP3;  
OS Apis mellifera (Honeybee).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Hymenoptera; Apoidea;  
OC Apidae; Apis.  
OX NCBI\_TaxID=7460;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=22448659; PubMed=12559578; DOI=10.1016/S0378-1119(02)01174-5;  
RA Malecova B., Ramser J., O'Brien J.K., Janitz M., Judova J.,  
RA Lebrach H., Simuth J.;  
RT "Honeybee (Apis mellifera L.) mrjp gene family: computational analysis  
of putative promoters and genomic structure of mrjp1, the gene coding  
for the most abundant protein of larval food.";  
RL Gene 303:165-175(2003).  
DR EMBL; AY078400; AAL83703.1; -; Genomic\_DNA.  
FT NON TER 31  
SQ SEQUENCE 31 AA; 3371 MW; 566F7E355EDC627B CRC64;

Query Match 100.0%; Score 49; DB 2; Length 31;  
Best Local Similarity 100.0%; Pred. No. 0.0055;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAVNHORKSA 10

DB 21 AAVNHORKSA 30

##### RESULT 2

MRJP3 APIME  
ID MRJP3 APIME STANDARD; PRT; 544 AA.  
AC Q17060;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-MAY-2005 (Rel. 47, Last annotation update)  
DE Major royal jelly protein 3 precursor (MRJP-3) (Bee-milk protein)  
DE (Royal jelly protein RJP57-1).  
GN Name=MRJP3;  
OS Apis mellifera (Honeybee).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Hymenoptera; Apoidea;  
OC Apidae; Apis.  
OX NCBI\_TaxID=7460;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RT TISSUE=Head;

RA Klaudiny J., Hanes J., Kulifajova J., Albert S., Simuth J.;  
RT "Molecular cloning of two cDNAs from the head of the nurse honey bee  
RL (*Apis mellifera* L.) for coding related proteins of royal jelly.";  
RN J. Apicultural Res. 33:105-111(1994).  
RP [2]  
RA SEQUENCE REVISION TO THE C-TERMINUS.  
RP Albert S., Klaudiny J., Simuth J.;  
RT "Newly discovered features of the updated sequence of royal jelly  
RT plectin RJP571; longer repetitive region on C-terminus and homology to  
RL *Drosophila melanogaster* yellow protein.";  
RN J. Apicultural Res. 35:63-68(1996).  
RA [43]  
RT NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 21-36.  
RC TISSUE=Head;  
RX MEDLINE=99007754; PubMed=9791542; DOI=10.1007/s000180050229;  
RA Schmitzova J., Klaudiny J., Albert S., Schroeder W., Schreckengost W.,  
RA Hanes J., Judova J., Simuth J.;  
RT "A family of major royal jelly proteins of the honeybee *Apis mellifera*  
RT L.";  
RL Cell. Mol. Life Sci. 54:1020-1030(1998).  
RN [4]  
RP NUCLEOTIDE SEQUENCE, AND PARTIAL PROTEIN SEQUENCE OF 45-59; 70-77;  
RP 106-120; 214-224 AND 243-253.  
RC TISSUE=Hypopharyngeal gland;  
RX MEDLINE=96055707; PubMed=9395329;  
RA Ohashi K., Natori S., Kubo T.;  
RT "Change in the mode of gene expression of the hypopharyngeal gland  
RT cells with an age-dependent role change of the worker honeybee *Apis*  
RT *mellifera* L.";  
RL Eur. J. Biochem. 249:797-802(1997).  
RN [4]  
RP -1- FUNCTION: May play an important role in honeybee nutrition. It is  
CC found in the royal jelly which is the food of the queen honey bee  
CC larva. The royal jelly determines the development of the young  
CC larvae and is responsible for the high reproductive ability of the  
CC honeybee queen.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- TISSUE SPECIFICITY: Hypopharyngeal glands of nurse honey bees.  
CC -1- DEVELOPMENTAL STAGE: Produced by the cephalic glandular system of  
CC the nurse honey bee.  
CC -1- SIMILARITY: Belongs to the major royal jelly protein family.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
DR EMBL; Z6318; CAA81227.1; -; mRNA.  
DR Ensembl; ENSGAPG00000007331; *Apis mellifera*.  
DR InterPro; IPR003534; Royaljelly.  
DR PANTHER; PTHR10009; Royaljelly; 1.  
DR Pfam; PF03022; MRJP; 1.  
DR PRINTS; PR01366; ROYALJELLY.  
RW Direct protein sequencing; Glycoprotein; Repeat; Signal.  
FT SIGNAL  
FT CHAIN 1 20  
FT REPEAT 21 544 Major royal jelly protein 3.  
FT REPEAT 424 428 1.  
FT REPEAT 429 433 2.  
FT REPEAT 434 438 3.  
FT REPEAT 439 443 4.  
FT REPEAT 444 448 5.  
FT REPEAT 449 453 6.  
FT REPEAT 454 458 7.  
FT REPEAT 459 463 8.  
FT REPEAT 464 468 9.  
FT REPEAT 469 473 10.  
FT REPEAT 474 478 11.  
FT REPEAT 479 483 12.  
FT REPEAT 484 488 13.  
FT REPEAT 489 493 14.  
FT REPEAT 494 498 15.  
FT REPEAT 499 503 16.  
FT REPEAT 504 508 17.

FT REPEAT 509 513 18.  
FT REPEAT 514 518 19.  
FT REPEAT 519 523 20.  
FT REGION 424 523 20 X 5 AA tandem repeats of [NKR]-Q-N-  
FT [AGD]-[DNG].  
FT CARBOHYD 183 183 N-linked (GlcNAc...) (Potential).  
SQ SEQUENCE 544 AA; 61662 MW; 4C8FFFC8A2759F52 CRC64;  
Query Match 100.0%; Score 49; DB 1; Length 544;  
Best Local Similarity 100.0%; Pred. No. 0.14;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AAVNHQKSA 10  
Db 21 AAVNHQKSA 30  
|||||  
RESULT 3  
Q7P382\_FUSNV  
ID Q7P382\_FUSNV PRELIMINARY; PRT; 344 AA.  
AC Q7P382;  
DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Hypothetical cytosolic protein.  
GN Name=FN2194;  
OS Fusobacterium nucleatum subsp. vincentii ATCC 49256.  
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;  
OC Fusobacterium.  
OX NCBI\_TaxID=209882;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=ATCC 49256;  
RA Karpatri V., Ivanova N., Anderson I., Reznik G., Bhattacharyya A.,  
RA Gardner W.L., Mikhailova N., Larsen N., D'Souza M., Walunas T.,  
RA Haselkorn R., Overbeek R., Kyriakides N.;  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
CC -1- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL; AABF01000005; EAA25090.1; -; Genomic\_DNA.  
DR InterPro; IPR009362; DUF1016.  
DR Pfam; PF06250; DUF1016; 1.  
RW Hypothetical protein.  
SQ SEQUENCE 344 AA; 40197 MW; B7117B1C125645D9 CRC64;  
Query Match 81.6%; Score 40; DB 2; Length 344;  
Best Local Similarity 87.5%; Pred. No. 7.1;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AAVNHQK 8  
Db 275 AAVNHQK 282  
|||||  
RESULT 4  
Q8UI59\_AGR75  
ID Q8UI59\_AGR75 PRELIMINARY; PRT; 169 AA.  
AC Q8UI59;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 24, Last annotation update)  
DE Hypothetical protein Atu0441.  
GN OrderedLocustNames=Atu0441;  
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.  
OX NCBI\_TaxID=176299;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=Dupont;  
RX MEDLINE=21608550; PubMed=11743193; DOI=10.1126/science.1066804;  
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,

RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,  
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,  
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,  
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,  
RA Raymond C., Rouge G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,  
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,  
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,  
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,  
RA Nester E.W.;  
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens  
RT C58.";  
RL Science 294:2317-2323(2001).  
DR ENBL: AE009014; AAL41460.1; -; Genomic\_DNA.  
DR PR: AP2630; AP2630.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 169 AA; 19428 MW; FDBBB0BAC5D38EB CRC64;  
  
Query Match 77.6%; Score 38; DB 2; Length 169;  
Best Local Similarity 70.0%; Pred. No. 8.6;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 AAVNHOKRSA 10  
Db :|||||:  
17 SAVNHSRSEA 26  
  
RESULT 5  
MOBA\_ECOLI  
ID MOBA\_ECOLI STANDARD; PRT; 194 AA.  
AC P58221;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Molybdopterin-guanine dinucleotide biosynthesis protein A (Protein  
DE FA).  
GN Name=moba; Synonyms=chlB, mob, narB; OrderedLocusNames=z5389, ECs4780;  
OS Escherichia coli O157:H7.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=83334;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;  
RX MEDLINE=21074935; PubMed=11205551; DOI=10.1038/35054089;  
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,  
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
RA Welch R.A., Blattner F.R.;  
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";  
RL Nature 409:529-533(2001).  
RN [2]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC;  
RX MEDLINE=21156231; PubMed=11258796;  
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,  
RA Kihara S., Shiba T., Hattori M., Shinagawa H.,  
RA "Complete genome sequence of enterohaemorrhagic Escherichia coli  
RT O157:H7 and genomic comparison with a laboratory strain K-12.";  
RL DNA Res. 8:11-22(2001).  
RN [3]  
RP NUCLEOTIDE SEQUENCE OF 87-194.  
RC STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC;  
RX Ohnishi M., Murata T., Nakayama K., Kuhara S., Hattori M.,  
RA Kurokawa K., Yasunaga T., Makino K., Shinagawa H., Hayashi T.;  
RT "Development of primer sets for direct sequence determination of all  
RT the ribosomal operons of Escherichia coli.";  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Links a guanosine 5'-phosphate to molybdopterin (MPT)  
CC forming molybdopterin guanine dinucleotide (MGD) (By similarity).

CC -!- COFACTOR: Magnesium (By similarity).  
CC -!- PATHWAY: Molybdenum cofactor biosynthesis.  
CC -!- SUBUNIT: Monomer (By similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: Belongs to the mobA family.  
CC  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC  
CC ENBL: AE005174; AAG59046.1; -; Genomic DNA.  
DR ENBL: BA000007; BAB38203.1; -; Genomic DNA.  
DR ENBL: AB035920; BAA93557.1; -; Genomic DNA.  
DR PR: D91226; D91226.  
DR HSP: P32173; 1FR9.  
DR SMR: P58221; 4-191.  
DR HAMAP: MF 00316; -; 1.  
KW Complete proteome; GTP-binding; Magnesium;  
KW Molybdenum cofactor biosynthesis; Nucleotide-binding.  
SQ SEQUENCE 194 AA; 21629 MW; E6F2ABDEC7BA2205 CRC64;  
  
Query Match 75.5%; Score 37; DB 1; Length 194;  
Best Local Similarity 70.0%; Pred. No. 17;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 AAVNHOKRSA 10  
Db :|||||:  
111 ARLNHOKRDA 120  
  
RESULT 6  
MOBA\_ECOLI  
ID MOBA\_ECOLI STANDARD; PRT; 194 AA.  
AC P32173; Q9LBV0;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Molybdopterin-guanine dinucleotide biosynthesis protein A (Protein  
DE FA).  
GN Name=moba; Synonyms=chlB, mob, narB; OrderedLocusNames=b3857;  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=K12 / MG1655;  
RX MEDLINE=93347969; PubMed=8346018;  
RA Plunkett G. III, Burland V., Daniels D.L., Blattner F.R.;  
RT "Analysis of the Escherichia coli genome. III. DNA sequence of the  
RT region from 87.2 to 89.2 minutes.";  
RL Nucleic Acids Res. 21:3391-3398(1993).  
RN [2]  
RP PROTEIN SEQUENCE OF 1-5, AND CHARACTERIZATION.  
RX MEDLINE=94291668; PubMed=8020507;  
RA Palmer T., Vasishta A., Whitty P.W., Boxer D.H.;  
RT "Isolation of protein FA, a product of the mob locus required for  
RT molybdenum cofactor biosynthesis in Escherichia coli.";  
RL Eur. J. Biochem. 222:687-692(1994).  
RN [3]  
RP CHARACTERIZATION.  
RC STRAIN=K12;  
RX MEDLINE=96004466; PubMed=7551035;  
RA Iobbi-Nivol C., Palmer T., Whitty P.W., McNairn E., Boxer D.H.;  
RT "The mob locus of Escherichia coli K12 required for molybdenum  
RT cofactor biosynthesis is expressed at very low levels.";  
RL Microbiology 141:1663-1671(1995).  
RN [4]  
RP X-RAY CRYSTALLOGRAPHY (1.65 ANGSTROMS).  
RX MEDLINE=20568278; PubMed=10978347; DOI=10.1074/jbc.M007406200;  
RA Lake M.W., Temple C.A., Rajagopalan K.V., Schindelin H.;

RT "The crystal structure of the Escherichia coli Moba protein provides insight into molybdopterin guanine dinucleotide biosynthesis.";

RL J. Biol. Chem. 275:40211-40217(2000).

RN [5]

RP X-RAY CRYSTALLOGRAPHY (1.35 ANGSTROMS).

RX PubMed=11080634; DOI=10.1016/S0969-2126(00)00518-9;

RA Stevenson C.E., Sargent F., Buchanan G., Palmer T., Lawson D.M.;

RA "Crystal structure of the molybdenum cofactor biosynthesis protein Moba from Escherichia coli at near-atomic resolution.";

RT Structure 8:1115-1125(2000).

RL CC -1- FUNCTION: Links a guanosine 5'-phosphate to molybdopterin (MPT) forming molybdopterin guanine dinucleotide (MGD).

CC CC -1- COFACTOR: Magnesium.

CC CC -1- PATHWAY: Molybdenum cofactor biosynthesis.

CC CC -1- SUBUNIT: Monomer.

CC CC -1- SUBCELLULAR LOCATION: Cytoplasmic.

CC CC -1- SIMILARITY: Belongs to the mobA family.

CC -----

CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

CC -----

DR EMBL; L19201; AAB02992.1; -; Genomic\_DNA.

DR EMBL; U00096; AAC76855.1; -; Genomic\_DNA.

DR PIR; S40803; S40803.

DR PDB; 1E5K; X-ray; A=1-194.

DR PDB; 1FR9; X-ray; A=1-194.

DR PDB; 1FRW; X-ray; A=1-194.

DR PDB; 1H4C; X-ray; A=1-194.

DR PDB; 1H4D; X-ray; A=1-194.

DR PDB; 1H4E; X-ray; A=1-194.

DR PDB; 1HJG; X-ray; A=1-194.

DR PDB; 1HJL; X-ray; A=1-194.

DR EchoBASE; EBI776; -.

DR EcoGene; EGI1829; mobA.

DR HAMAP; MF 00316; -; 1.

DR 3D-structure; Complete proteome; Direct protein sequencing;

KW GTP-binding; Magnesium; Molybdenum cofactor biosynthesis;

KW Nucleotide-binding.

FT STRAND 7 12

FT HELIX 25 27

FT STRAND 29 30

FT TURN 31 32

FT STRAND 33 34

FT HELIX 35 46

FT STRAND 50 53

FT HELIX 58 62

FT TURN 63 64

FT STRAND 68 69

FT TURN 73 74

FT TURN 79 89

FT HELIX 94 99

FT TURN 100 101

FT TURN 103 104

FT TURN 107 108

FT HELIX 109 115

FT TURN 116 116

FT TURN 118 119

FT STRAND 122 126

FT STRAND 131 139

FT TURN 140 141

FT HELIX 142 151

FT TURN 152 153

FT HELIX 157 163

FT TURN 164 165

FT STRAND 167 170

FT TURN 172 173

FT TURN 175 178

FT HELIX 184 188

FT TURN 189 190

FT SEQUENCE 194 AA; 21643 MW; B79B32DD7348DD48 CRC64;

Query Match 75.5%; Score 37; DB 1; Length 194;  
Best Local Similarity 70.0%; Pred. No. 17;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAVNHORKSA 10  
DB 111 ARLNHORKDA 120

## RESULT 7

Q83PF7\_SHIFL Q83PF7\_SHIFL PRELIMINARY; PRT; 194 AA.  
AC Q83PF7; Q7BZD5;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DE Molybdopterin-guanine dinucleotide biosynthesis protein A.  
GN Names-moba; OrderedLocustNames=S3819, SF3928;  
OS Shigella flexneri.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Shigella.  
OX NCBI\_TaxID=623;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=301 / Serotype 2a;  
RX MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;  
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,  
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,  
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,  
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,  
RA Yu J.;  
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity through comparison with genomes of Escherichia coli K12 and O157.";  
RL Nucleic Acids Res. 30:4432-4441(2002).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;  
RX MEDLINE=22590274; PubMed=12704152;  
RX DOI=10.1128/IAI.71.5.2775-2786.2003;  
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,  
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,  
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,  
RA Schwartz D.C., Blattner F.R.;  
RT "Complete genome sequence and comparative genomics of Shigella flexneri serotype 2a strain 2457T.";  
RL Infect. Immun. 71:2775-2786(2003).  
DR EMBL; AE005674; AAN45363.1; -; Genomic DNA.  
DR EMBL; AE016990; AAP18835.1; -; Genomic DNA.  
DR HSSP; P32173; 1FR9.  
DR SMR; Q83PF7; 4-191.  
KW Complete proteome.  
SQ SEQUENCE 194 AA; 21608 MW; 8933CCDECDDB46BE CRC64;

Query Match 75.5%; Score 37; DB 2; Length 194;  
Best Local Similarity 70.0%; Pred. No. 17;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAVNHORKSA 10  
DB 111 ARLNHORKDA 120

## RESULT 8

Q8BN36\_MOUSE Q8BN36\_MOUSE PRELIMINARY; PRT; 135 AA.  
AC Q8BN36;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Mus musculus activated spleen cDNA, RIKEN full-length enriched DE library, clone:F830008E12 product:hypothetical protein, full insert DE sequence.

OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=NOD; TISSUE=Activated spleen;  
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
RA Carninci P., Hayashizaki Y.;  
RT "High-efficiency full-length cDNA cloning.";  
RL Meth. Enzymol. 303:19-44(1999).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=NOD; TISSUE=Activated spleen;  
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staudt P., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Oikido T., Furuno M., Bono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=NOD; TISSUE=Activated spleen;  
RA The PANTOM Consortium,  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.";  
RL Nature 420:563-573(2002).  
RN [4]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=NOD; TISSUE=Activated spleen;  
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
RL Genome Res. 10:1617-1630(2000).  
RN [5]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=NOD; TISSUE=Activated spleen;  
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,  
RA Sui N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
RT "RIKEN integrated sequence analysis (RISA) system-384-format  
RT sequencing pipeline with 384 multicapillary sequencer.";  
RL Genome Res. 10:1757-1771(2000).  
RN [6]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=NOD; TISSUE=Activated spleen;  
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,

RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,  
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Iegami M.,  
RA Tagawa A., Takahashi P., Takaku-Akahira S., Takeda Y., Tanaka T.,  
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR ENBL; AK089672; BC40942.1; -; mRNA.  
DR Ensembl; ENSMUSG00000063962; Mus musculus.  
KW Hypothetical protein.  
SQ SEQUENCE 135 AA; 15774 MW; 5EP998816A8248EB CRC64;  
  
Query Match 73.5%; Score 36; DB 2; Length 135;  
Best Local Similarity 87.5%; Pred. No. 18;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 AVNHQRKS 9  
DB 88 AVHHQRKS 95  
||:|||||  
||:|||||  
  
RESULT 9  
O97459 DROME PRELIMINARY; PRT; 347 AA.  
AC O97459;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)  
DE HCG-1.  
GN Name=UBL3; Synonyms=HCG-1; ORFNames=CG9038;  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=99307157; PubMed=10375635; DOI=10.1016/S0378-1119(99)00138-9;  
RA Chadwick B.P., Kidd T., Sgouros J., Ish-Horowitz D., Frischaut A.-M.;  
RT "Cloning, mapping and expression of UBL3, a novel ubiquitin-like  
RT gene.";  
RL Gene 233:189-195(1999).  
DR ENBL; AF044220; AAD02322.1; -; Genomic\_DNA.  
DR ENBL; AF044219; AAD02321.1; -; mRNA.  
DR SMR; O97459; 231-335.  
DR Ensembl; CG9038; Drosophila melanogaster.  
DR FlyBase; FBgn0026076; CG9038.  
DR FlyBase; FBgn0026076; UBL3.  
DR GO; GO:0006464; P:protein modification; IEA.  
DR InterPro; IPR000626; Ubiquitin.  
DR PROSITE; PS0053; UBIQUITIN\_2; 1.  
SQ SEQUENCE 347 AA; 38086 MW; 5AC852F01D7AEFA7 CRC64;  
  
Query Match 73.5%; Score 36; DB 2; Length 347;  
Best Local Similarity 77.8%; Pred. No. 52;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 AAVNHQRKS 9  
DB 145 AAVHHQRKS 153  
||:|||||  
||:|||||  
  
RESULT 10  
Q9VXN1 DROME  
ID Q9VXN1 DROME PRELIMINARY; PRT; 347 AA.  
AC Q9VXN1;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)  
DE CG9038-PA, isoform A (CG9038-pb, isoform b) (LD03728p).  
GN Name=UBL3; ORFNames=CG9038,

OS Drosophila melanogaster (fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN NUCLEOTIDE SEQUENCE.  
 RP MEDLINE=2019606; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RX Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Anantides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazie R.G., Champe M., Pfeiffer B.D.,  
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,  
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Eosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D.A., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laeko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler P., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*."  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=22426065; PubMed=12537568;  
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
 RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,  
 RA Pacieb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstein G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
 RT "Finishing a whole-genome shotgun: release 3 of the *Drosophila*  
 RT melanogaster euchromatic genome sequence."  
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,  
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celniker S.E.;  
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin:  
 RT a genome perspective."  
 RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).  
 RN [4]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Miera S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,

RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review."  
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
 RN [5]  
 RP NUCLEOTIDE SEQUENCE.  
 RG Berkeley *Drosophila* Genome Project;  
 RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,  
 RA Hoskins R., Stapleton M., Pacieb J., Park S., Svirskas R., Smith E.,  
 RA Yu C., Rubin G.;  
 RT "Drosophila melanogaster release 4 sequence."  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP NUCLEOTIDE SEQUENCE.  
 RG FlyBase;  
 RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=Berkley;  
 RA Stapleton M., Carlson J., Chavez C., Frise E., George R., Pacieb J.,  
 RA Park S., Wan K., Yu C., Rubin G.M., Celniker S.;  
 RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.  
 CC -I- INTERACTION:  
 CC Q9VSV3:CG4483; NBExp=1; InAct=EBI-165186, EBI-176779;  
 CC EMBL; AE003500; AAF48527.1; -; Genomic\_DNA.  
 CC EMBL; BT021420; AAX33568.1; -; mRNA.  
 DR SNR; Q9VKN1; 231-335.  
 DR IntAct; Q9VKN1; -;  
 DR Ensembl; CG9038; *Drosophila melanogaster*.  
 DR FlyBase; FBgn0026076; CG9038.  
 DR FlyBase; FBgn0026076; UBL3.  
 DR GO; GO:006464; P:protein modification; IEA.  
 DR InterPro; IPR000626; Ubiquitin.  
 DR PROSITE; PS50053; UBIQUITIN\_2; 1.  
 DR SEQUENCE 347 AA; 38072 MW; AF9D07A71D7AEFA6 CRC64;  
 SQ  
 Query Match 73.5%; Score 36; DB 2; Length 347;  
 Best Local Similarity 77.8%; Pred. No. 52;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AAVNHQKS 9  
 DB 145 AAVHHQQS 153  
 |||:||||:  
 |||:||||:  
 RESULT 11  
 Q8SX00 DROME PRELIMINARY; PRT; 375 AA.  
 ID Q8SX00 DROME PRELIMINARY; PRT; 375 AA.  
 AC Q8SX00;  
 DT 01-JUN-2002 (T-EMBLrel. 21, Created)  
 DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)  
 DT 01-OCT-2002 (T-EMBLrel. 22, Last annotation update)  
 DE RH47633P.  
 GN ORFNames=CG8311;  
 OS *Drosophila melanogaster* (fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=Berkley;  
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
 RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,  
 RA Miranda A., Mungall C.J., Nunoo J., Pacieb J., Faragas V., Park S.,  
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
 RA Celniker S.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; AY094927; AM11280.1; -; mRNA.

DR Ensembl; CG8311; Drosophila melanogaster.  
 DR FlyBase; FBgn0034141; CG8311.  
 SQ SEQUENCE 375 AA; 41503 MW; B1212E3AF5F79ED5 CRC64;  
 Query Match 73.5%; Score 36; DB 2; Length 375;  
 Best Local Similarity 77.8%; Pred. No. 57;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Qy 2 AVNHQKSA 10  
 Db 29 AVRQRRSA 37  
 RESULT 12  
 Q9V7R7 DROME PRELIMINARY; PRT; 503 AA.  
 AC Q9V7R7;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE CG8311-PA.  
 GN Name=CG8311; ORFNames=CG8311;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burtis K.C., Buesam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K.J., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.A., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laoko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle B.J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupek M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195 (2000).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=22426065; PubMed=12537573;  
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,

George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,  
 Pacle B.J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstock G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
 RT "Finishing a whole-genome shotgun: release 3 of the Drosophila  
 melanogaster euchromatic genome sequence."  
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,  
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celniker S.E.;  
 RT "The transposable elements of the Drosophila melanogaster euchromatin:  
 a genomic perspective."  
 RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).  
 RN [4]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Battencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
 systematic review."  
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
 RN [5]  
 RP NUCLEOTIDE SEQUENCE.  
 RG Berkeley Drosophila Genome Project;  
 RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,  
 RA Hoskins R., Stapleton M., Pacle B.J., Park S., Svirskas R., Smith E.,  
 RA Yu C., Rubin G.;  
 RT "Drosophila melanogaster release 4 sequence."  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP NUCLEOTIDE SEQUENCE.  
 RG FlyBase;  
 RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE003806; ANF57978.1; -; Genomic DNA.  
 DR Ensembl; CG8311; Drosophila melanogaster.  
 DR FlyBase; FBgn0034141; CG8311.  
 DR GO; GO:0016020; C:nembrane; IEA.  
 DR GO; GO:0004605; P:phosphatidate cytidyltransferase activity; IEA.  
 DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.  
 DR InterPro; IPR000374; PC trans.  
 DR Pfam; PF01148; CTP transf. 1;  
 SQ SEQUENCE 503 AA; 54801 MW; 0179BA78BF307FB9 CRC64;  
 Query Match 73.5%; Score 36; DB 2; Length 503;  
 Best Local Similarity 77.8%; Pred. No. 79;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Qy 2 AVNHQKSA 10  
 Db 29 AVRQRRSA 37  
 RESULT 13  
 Q7NAK1 MYCGA PRELIMINARY; PRT; 277 AA.  
 ID Q7NAK1 MYCGA PRELIMINARY; PRT; 277 AA.  
 AC Q7NAK1;  
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Predicted transposase.  
 GN OrderedLocustNames=MYCGA6490; ORFNames=MGA\_0475;  
 OS Mycoplasma gallisepticum.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2096;  
 RN [1]



```
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=R(1ow);
RX MEDLINE=22830409; PubMed=12949158; DOI=10.1099/mic.0.26427-0;
RA Papazisi L., Gorton T.S., Kutish G., Markham P.F., Browning G.F.,
RA Nguyen D.K., Swartzell S., Madan A., Mahairas G., Geary S.J.;
RT "The complete genome sequence of the avian pathogen Mycoplasma
RT gallisepticum strain R(1ow).";
RL Microbiology 149:2307-2316 (2003).
DR EMBL; AE016969; AAP56999.1; -; Genomic_DNA.
DR InterPro; IPR001611; LRR.
DR PRINTS; PR00019; LEURICRPT.
KW Complete proteome; Leucine-rich repeat; Repeat.
SQ SEQUENCE 277 AA; 32757 MW; 9D52D025DBAC1BC5 CRC64;

Query Match 71.4%; Score 35; DB 2; Length 277;
Best Local Similarity 75.0%; Pred. No. 66;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 AVNHQKSA 9
Db 204 AINHONKS 211
|:|||||

RESULT 14
Q4HFA2_CAMCO PRELIMINARY; PRT; 304 AA.
AC Q4HFA2;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hemolysin, putative.
GN ORFNames=CC00183;
OS Campylobacter coli RM2228.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=306254;
RN 11
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RM2228;
RA Fouts D.E., Mongodin E.F., Mandrell R.E., Miller W.G., Rasko D.A.,
RA Jacques R.J., Brinkac L.M., DeBoy R.T., Parker C.T., Daugherty S.C.,
RA Dodson R.J., Durkin A.S., Madupu R.R., Sullivan S.A., Shetty J.U.,
RA Ayodeji M.A., Shvartsbeyn A.A., Schatz M.C., Badger J.H., Fraser C.M.,
RA Nelson K.E.;
RT "Major structural and novel potential virulence mechanisms from the
RT genomes of multiple Campylobacter species.";
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAPL0100006; EAL56539.1; -; Genomic DNA.
SQ SEQUENCE 304 AA; 34186 MW; 9BB9B7315D26FE51 CRC64;

Query Match 71.4%; Score 35; DB 2; Length 304;
Best Local Similarity 77.8%; Pred. No. 74;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 AVNHQKSA 10
Db 117 AKNHQESA 125
|:|||||

RESULT 15
Q6MIG1_BDEBA PRELIMINARY; PRT; 357 AA.
ID Q6MIG1_BDEBA PRELIMINARY;
AC Q6MIG1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE UDP-N-acetylglucosamine-N-acetylmuramyl-(Pentapeptide)
DE pyrophosphoryl-undecaprenol N-acetylglucosamine transferase
DE (EC 2.4.1.227).
GN Name=murG; OrderedLocusNames=Bd3197;
```

```
OS Bdellovibrio bacteriovorus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;
OC Bdellovibrionaceae; Bdellovibrio.
OX NCBI_TaxID=959;
RN 11
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HD100 / DSM 50701 / ATCC 15356 / NCIB 9529;
RX PubMed=14752164; DOI=10.1126/science.1093027;
RA Rendulic S., Jagtap P., Rosinus A., Eppinger M., Baar C., Lanz C.,
RA Keller H., Lambert C., Evans K.J., Gossman A., Meyer F.,
RA Sockett R.E., Schuster S.C.;
RT "A predator unmasked: life cycle of Bdellovibrio bacteriovorus from a
RT genomic perspective.";
RL Science 303:689-692 (2004).
DR EMBL; BX842654; CAE80952.1; -; Genomic DNA.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0030246; F:carbohydrate binding; IEA.
DR GO; GO:0016758; F:transferase activity, transferring hexosyl . . . ; IEA.
DR GO; GO:0050511; F:undecaprenyldiphospho-muramoylpentapeptide . . . ; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0030259; P:lipid glycosylation; IEA.
DR GO; GO:0019277; P:UDP-N-acetylgalactosamine biosynthesis; IEA.
DR InterPro; IPR004276; Glyco_trans_28.
DR InterPro; IPR007235; Glyco_tran_28_C.
DR InterPro; IPR006009; MurG.
DR Pfam; PF03033; Glyco_transf_28; 1.
DR Pfam; PF04101; Glyco_tran_28_C; 1.
DR TIGRFAMs; TIGR01133; murG; 1.
DR Complete proteome; Glycosyltransferase; Transferase.
KW SEQUENCE 357 AA; 38880 MW; 64C993ACEA94D881 CRC64;

Query Match 71.4%; Score 35; DB 2; Length 357;
Best Local Similarity 70.0%; Pred. No. 88;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAVNHQKSA 10
Db 287 AADNHQKNA 296
|:|||||

Search completed: May 3, 2006, 19:08:58
Job time : 8.62963 secs
```



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2006, 19:09:24 ; Search time 1.08025 Seconds  
(without alignments)  
765.341 Million cell updates/sec

Title: US-10-525-567-1

Perfect score: 49

Sequence: 1 AAVNHQKSA 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/1/iaa/5 COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/6 COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/H COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/PCTRUS COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/RE COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	36	73.5	170	2 US-09-270-767-44218
2	36	73.5	423	2 US-09-270-767-41561
3	33	67.3	355	2 US-09-248-796A-14612
4	32	65.3	77	2 US-09-583-110-4552
5	32	65.3	242	2 US-09-902-540-13487
6	32	65.3	249	2 US-09-107-532A-6748
7	32	65.3	302	2 US-09-252-991A-28202
8	32	65.3	309	2 US-09-489-039A-9949
9	32	65.3	504	1 US-08-896-005-3
10	32	65.3	692	2 US-09-487-538B-274
11	31	63.3	26	1 US-08-620-151-28
12	31	63.3	38	2 US-09-774-639-254
13	31	63.3	68	2 US-09-248-796A-24682
14	31	63.3	108	2 US-09-489-039A-10310
15	31	63.3	185	2 US-09-949-016-8153
16	31	63.3	221	2 US-09-902-540-14900
17	31	63.3	361	2 US-09-252-991A-30237
18	31	63.3	365	2 US-09-252-991A-20893
19	31	63.3	407	2 US-09-252-991A-20503
20	31	63.3	419	2 US-09-902-540-16674
21	31	63.3	432	2 US-09-252-991A-23622
22	31	63.3	433	1 US-08-417-492-2
23	31	63.3	458	2 US-09-328-352-5083
24	31	63.3	467	2 US-09-489-039A-10022
25	31	63.3	569	2 US-09-902-540-11992
26	31	63.3	667	2 US-09-252-991A-32588
27	31	63.3	1307	2 US-09-949-016-7561

28	31	63.3	2440	2	US-09-632-033B-3	Sequence 3, Appli
29	31	63.3	2451	2	US-09-949-016-9675	Sequence 9675, Ap
30	30	61.2	65	2	US-09-107-433-2735	Sequence 2735, Ap
31	30	61.2	71	2	US-09-270-767-57778	Sequence 57778, A
32	30	61.2	137	2	US-09-134-000C-5507	Sequence 5507, Ap
33	30	61.2	162	2	US-09-270-767-57842	Sequence 57842, A
34	30	61.2	185	2	US-09-270-767-42540	Sequence 42540, A
35	30	61.2	202	2	US-09-252-991A-29151	Sequence 29151, A
36	30	61.2	225	2	US-09-252-991A-25020	Sequence 25020, A
37	30	61.2	221	2	US-09-605-703B-1648	Sequence 1648, Ap
38	30	61.2	276	1	US-07-952-817-27	Sequence 27, Appl
39	30	61.2	359	2	US-09-270-767-40759	Sequence 40759, A
40	30	61.2	359	2	US-09-270-767-55975	Sequence 55975, A
41	30	61.2	481	2	US-09-489-039A-8415	Sequence 8415, Ap
42	30	61.2	483	2	US-08-369-822C-22	Sequence 22, Appl
43	30	61.2	483	2	US-08-582-776C-37	Sequence 37, Appl
44	30	61.2	483	2	US-08-434-831B-34	Sequence 34, Appl
45	30	61.2	543	2	US-09-902-540-16697	Sequence 16697, A

## ALIGNMENTS

### RESULT 1

US-09-270-767-44218  
; Sequence 44218, Application US/09270767  
; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 44218  
; LENGTH: 170

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

US-09-270-767-44218

Query Match 73.5%; Score 36; DB 2; Length 170;

Best Local Similarity 77.8%; Pred. No. 9.7;

Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAVNHQKRS 9

Db 74 AAVHHQKRS 82

### RESULT 2

US-09-270-767-41561  
; Sequence 41561, Application US/09270767  
; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 41561  
; LENGTH: 423

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

; FEATURE:

; OTHER INFORMATION: Xaa means any amino acid

US-09-270-767-41561

Query Match 73.5%; Score 36; DB 2; Length 423;

Best Local Similarity 77.8%; Pred. No. 27;

Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 AAVNHQKRS 9
Db 221 AAVHQHQRQS 229

RESULT 3
US-09-248-796A-14612
; Sequence 14612, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 14612
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-14612

Query Match 67.3%; Score 33; DB 2; Length 355;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 NHQKRS 9
Db 279 NHQKRS 284

RESULT 4
US-09-583-110-4552
; Sequence 4552, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 4552
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-4552

Query Match 65.3%; Score 32; DB 2; Length 77;
Best Local Similarity 60.0%; Pred. No. 26;
Matches 6; Conservative 3; Mismatches 1; Indels 1; Gaps 0;

Qy 1 AAVNHQKSA 10
Db 10 AALAHQRRAA 19

RESULT 5
US-09-902-540-13487
; Sequence 13487, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 13487
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-13487

Query Match 65.3%; Score 32; DB 2; Length 242;
Best Local Similarity 75.0%; Pred. No. 93;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AVNHQKRS 9
Db 23 AVDHQKRS 30

RESULT 6
US-09-107-532A-6748
; Sequence 6748, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6748:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 249 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
```

```

;
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...249
; SEQUENCE DESCRIPTION: SEQ ID NO: 6748:
US-09-107-532A-6748
    Query Match      65.3%; Score 32; DB 2; Length 249;
    Best Local Similarity 75.0%; Pred. No. 96;
    Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      2 AVNHQKRS 9
Db      157 AVNHPRKN 164
      |||||
      |||||

RESULT 7
US-09-252-991A-28202
; Sequence 28202, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28202
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28202
    Query Match      65.3%; Score 32; DB 2; Length 302;
    Best Local Similarity 77.8%; Pred. No. 1.2e+02;
    Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 AAVNHQKRS 9
Db      41 AAVSHQVKS 49
      |||||
      |||||

RESULT 8
US-09-489-039A-9949
; Sequence 9949, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9949
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9949
    Query Match      65.3%; Score 32; DB 2; Length 309;
    Best Local Similarity 77.8%; Pred. No. 1.2e+02;
    Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 AAVNHQKRS 9
Db      39 AAVSHQVKS 47
      |||||
      |||||

FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...249
SEQUENCE DESCRIPTION: SEQ ID NO: 6748:
US-09-107-532A-6748
    Query Match      65.3%; Score 32; DB 2; Length 249;
    Best Local Similarity 75.0%; Pred. No. 96;
    Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      2 AVNHQKRS 9
Db      157 AVNHPRKN 164
      |||||
      |||||

RESULT 9
US-08-896-005-3
; Sequence 3, Application US/08896005
; Patent No. 5854023
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: S-ADENOSYL-L-HOMOCYSTEINE HYDROLASE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/896,005
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0337 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 969078
; US-08-896-005-3
    Query Match      65.3%; Score 32; DB 1; Length 504;
    Best Local Similarity 85.7%; Pred. No. 2.1e+02;
    Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      4 NHQKSA 10
Db      57 NHQRNSA 63
      |||||
      |||||

RESULT 10
US-09-487-558B-274
; Sequence 274, Application US/09487558B
; Patent No. 6949356
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. 6949356man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
```

;; APPLICANT: Sherman, Amir  
;; APPLICANT: Silva, Jeff  
;; APPLICANT: Summers, Eric  
;; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi  
;; FILE REFERENCE: 109272.130  
;; CURRENT APPLICATION NUMBER: US/09/487,558B  
;; CURRENT FILING DATE: 2000-01-19  
;; PRIOR APPLICATION NUMBER: US 60/487,558  
;; PRIOR FILING DATE: 1999-10-20  
;; NUMBER OF SEQ ID NOS: 446  
;; SOFTWARE: PatentIn version 3.0  
;; SEQ ID NO 274  
;; LENGTH: 692  
;; TYPE: PRT  
;; ORGANISM: Saccharomyces cerevisiae  
US-09-487-558B-274

Query Match 55.3%; Score 32; DB 2; Length 692;  
Best Local Similarity 75.0%; Pred. No. 3e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VNHOKSA 10  
| | | | |  
DB 75 VNHQHS 82

RESULT 11  
US-08-620-151-28  
; Sequence 28, Application US/08620151  
; Patent No. 5928955  
; GENERAL INFORMATION:  
; APPLICANT: Imperiali, Barbara  
; APPLICANT: Walkup, Grant K.  
; TITLE OF INVENTION: PEPTIDYL FLUORESCENT CHEMOSENSOR FOR  
; TITLE OF INVENTION: DIVALENT ZINC  
; NUMBER OF SEQUENCES: 136  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE  
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront  
; STREET: Plaza Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60611-5599  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/620,151  
; FILING DATE: 22-MAR-1996  
; CLASSIFICATION: 422  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Shannon, Karen L.  
; REGISTRATION NUMBER: 36,675  
; REFERENCE/DOCKET NUMBER: 8597/6  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-321-4200  
; TELEFAX: 312-321-4299  
; INFORMATION FOR SEQ ID NO: 28:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 26 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-620-151-28

Query Match 63.3%; Score 31; DB 1; Length 26;  
Best Local Similarity 85.7%; Pred. No. 13;  
Matches 6; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

QY 1 AAVNHOR 7  
| | | | |  
DB 16 ALVNHOR 22

RESULT 12  
US-09-774-639-254  
; Sequence 254, Application US/09774639  
; Patent No. 6806351  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 90 Human Secreted Proteins  
; FILE REFERENCE: P2013P1  
; CURRENT APPLICATION NUMBER: US/09/774,639  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/244,112  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-02-04  
; NUMBER OF SEQ ID NOS: 371  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 254  
; LENGTH: 38  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-774-639-254

Query Match 63.3%; Score 31; DB 2; Length 38;  
Best Local Similarity 62.5%; Pred. No. 19;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VNHOKSA 10  
: | : | : | : |  
DB 1 IRHERKSA 8

RESULT 13  
US-09-248-796A-24682  
; Sequence 24682, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS  
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 24682  
; LENGTH: 68  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-09-248-796A-24682

Query Match 63.3%; Score 31; DB 2; Length 68;  
Best Local Similarity 66.7%; Pred. No. 37;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 AVNHOKSA 10  
| | | | : |  
DB 35 APNHOKKA 43

RESULT 14  
US-09-489-039A-10310  
; Sequence 10310, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709,2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 10310  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-10310

Query Match 63.3%; Score 31; DB 2; Length 108;  
Best Local Similarity 70.0%; Pred.No. 61;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAVNHQKSA 10  
Db 98 AAVNDERPSA 107

RESULT 15  
US-09-949-016-8153  
; Sequence 8153, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 8153  
; LENGTH: 185  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-8153

Query Match 63.3%; Score 31; DB 2; Length 185;  
Best Local Similarity 75.0%; Pred.No. 1.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VNHQKSA 10  
Db 68 LNHVRKSA 75

Search completed: May 3, 2006, 19:12:04  
Job time : 2.08025 secs

end Page Blank (uspro)

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2006, 19:10:39 ; Search time 3.50823 Seconds  
(without alignments)  
1190.997 Million cell updates/sec

Title: US-10-525-567-1  
Perfect score: 49  
Sequence: 1 AAVNHQRKSA 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA Main.\*  
1: /cgm2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pap.\*  
2: /cgm2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pap.\*  
3: /cgm2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pap.\*  
4: /cgm2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pap.\*  
5: /cgm2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pap.\*  
6: /cgm2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	38	77.6	186	4	US-10-437-963-119181
2	36	73.5	347	6	US-11-097-143-19335
3	36	73.5	503	6	US-11-097-143-18111
4	36	73.5	599	5	US-10-450-763-44531
5	35	71.4	664	4	US-10-225-066A-1084
6	35	71.4	664	4	US-10-374-780A-2530
7	35	71.4	664	5	US-10-225-066A-1084
8	34	69.4	284	4	US-10-369-493-19873
9	34	69.4	291	5	US-10-450-763-38089
10	34	69.4	300	5	US-10-450-763-44774
11	34	69.4	445	5	US-10-450-763-35911
12	34	69.4	453	5	US-10-450-763-44530
13	34	69.4	454	5	US-10-450-763-36492
14	33	67.3	57	4	US-10-425-115-201509
15	33	67.3	108	4	US-10-425-115-453427
16	33	67.3	153	4	US-10-424-599-256805
17	33	67.3	250	4	US-10-282-122A-48895
18	33	67.3	291	5	US-10-450-763-38742
19	33	67.3	375	6	US-11-097-143-30918
20	33	67.3	1093	4	US-10-437-963-167806
21	33	67.3	3572	4	US-10-471-450-16
22	32	65.3	57	4	US-10-425-115-271519
23	32	65.3	58	4	US-10-425-115-287460
24	32	65.3	69	4	US-10-425-115-189718
25	32	65.3	71	5	US-10-450-763-33483
26	32	65.3	75	4	US-10-424-599-279595
27	32	65.3	83	5	US-10-450-763-56761

28	65.3	95	4	US-10-425-115-298869	Sequence 298869,
29	65.3	99	4	US-10-437-963-151148	Sequence 151148,
30	65.3	106	4	US-10-425-115-259425	Sequence 259425,
31	65.3	114	4	US-10-424-599-194647	Sequence 194647,
32	65.3	136	4	US-10-424-599-229751	Sequence 229751,
33	65.3	196	4	US-10-425-115-358696	Sequence 358696,
34	65.3	267	4	US-10-425-115-321509	Sequence 321509,
35	65.3	282	4	US-10-437-963-148332	Sequence 148332,
36	65.3	297	4	US-10-425-115-348894	Sequence 348894,
37	65.3	305	3	US-09-815-242-10277	Sequence 10277, A
38	65.3	305	3	US-09-815-242-13882	Sequence 13882, A
39	65.3	317	4	US-10-282-122A-71687	Sequence 71687, A
40	65.3	395	4	US-10-437-963-194504	Sequence 194504,
41	65.3	400	4	US-10-156-761-10393	Sequence 10393, A
42	65.3	425	4	US-10-425-114-54716	Sequence 54716, A
43	65.3	502	6	US-11-097-143-34098	Sequence 34098, A
44	65.3	504	6	US-11-097-143-41922	Sequence 41922, A
45	65.3	563	4	US-10-424-599-162116	Sequence 162116,

## ALIGNMENTS

RESULT 1  
US-10-437-963-119181  
; Sequence 119181, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Bouharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 119181  
; LENGTH: 186  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(186)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_22422C.1.pap  
US-10-437-963-119181

Query Match 77.6%; Score 38; DB 4; Length 186;  
Best Local Similarity 70.0%; Pred. No. 18;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 AAVNHQRKSA 10  
| | | | | : | | |  
Db 160 AAVNHEQKKA 169

RESULT 2  
US-11-097-143-19335  
; Sequence 19335, Application US/11097143  
; Publication No. US20050208558A1  
; GENERAL INFORMATION:  
; APPLICANT: Venter, J. Craig  
; APPLICANT: et al.  
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
; TITLE OF INVENTION: DROSOPHILA GENES.

FILE REFERENCE: CL000728  
; CURRENT APPLICATION NUMBER: US/11/097,143  
; CURRENT FILING DATE: 2005-04-04  
; PRIOR APPLICATION NUMBER: 60/157,832  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: 60/160,191  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR APPLICATION NUMBER: 60/161,932  
; PRIOR FILING DATE: 1999-10-28  
; PRIOR APPLICATION NUMBER: 60/164,769  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR APPLICATION NUMBER: 60/173,383  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: 60/175,693  
; PRIOR FILING DATE: 2000-01-12  
; PRIOR APPLICATION NUMBER: 60/184,831  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/191,637  
; PRIOR FILING DATE: 2000-03-23  
; NUMBER OF SEQ ID NOS: 43008  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 19335  
; LENGTH: 347  
; TYPE: PRT  
; ORGANISM: DROSOPHILA  
US-11-097-143-19335

Query Match 73.5%; Score 36; DB 6; Length 347;  
Best Local Similarity 77.8%; Pred. No. 87;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAVNHQKRS 9  
Db 145 AAVHQRQS 153  
|||||:

## RESULT 3

US-11-097-143-18111  
; Sequence 18111, Application US/11097143  
; Publication No. US20050208558A1  
; GENERAL INFORMATION:  
; APPLICANT: Venter, J. Craig  
; APPLICANT: et al.  
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
; TITLE OF INVENTION: DROSOPHILA GENES.  
; FILE REFERENCE: CL000728  
; CURRENT APPLICATION NUMBER: US/11/097,143  
; CURRENT FILING DATE: 2005-04-04  
; PRIOR APPLICATION NUMBER: 60/157,832  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: 60/160,191  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR APPLICATION NUMBER: 60/161,932  
; PRIOR FILING DATE: 1999-10-28  
; PRIOR APPLICATION NUMBER: 60/164,769  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR APPLICATION NUMBER: 60/173,383  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: 60/175,693  
; PRIOR FILING DATE: 2000-01-12  
; PRIOR APPLICATION NUMBER: 60/184,831  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/191,637  
; PRIOR FILING DATE: 2000-03-23  
; NUMBER OF SEQ ID NOS: 43008  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 18111  
; LENGTH: 503  
; TYPE: PRT  
; ORGANISM: DROSOPHILA  
US-11-097-143-18111

Query Match 73.5%; Score 36; DB 6; Length 503;  
Best Local Similarity 77.8%; Pred. No. 1.3e+02;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 AVNHQRKSA 10  
Db 29 AVHQRSA 37  
|||||:

## RESULT 4

US-10-450-763-44531  
; Sequence 44531, Application US/10450763  
; Publication No. US20050196754A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 44531  
; LENGTH: 599  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)...(599)  
; OTHER INFORMATION: Xaa = X or \* as defined in Table 2  
US-10-450-763-44531

Query Match 73.5%; Score 36; DB 5; Length 599;  
Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAVNHQK 8  
Db 141 AAINHQK 148  
|||||:

## RESULT 5

US-10-225-066A-1084  
; Sequence 1084, Application US/10225066A  
; Publication No. US20030226173A1  
; GENERAL INFORMATION:  
; APPLICANT: Mendel Biotechnology, Inc.  
; APPLICANT: RATCLIFFE, Oliver  
; APPLICANT: RIECHMANN, Jose Luis  
; APPLICANT: ADAM, Luc J  
; APPLICANT: DUBELL, Arnold T  
; APPLICANT: HEARD, Jacqueline E  
; APPLICANT: PILGRIM, Marsha L  
; APPLICANT: JIANG, Cai-Zhong  
; APPLICANT: REUBER, T. Lynne  
; APPLICANT: CREELMAN, Robert A  
; APPLICANT: PINEDA, Omalra  
; APPLICANT: YU, Guo-Liang  
; APPLICANT: BROUN, Pierre E  
; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants  
; FILE REFERENCE: MBI0036-2 US  
; CURRENT APPLICATION NUMBER: US/10/225,066A  
; CURRENT FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: 09/837,444  
; PRIOR FILING DATE: 2001-04-18  
; PRIOR APPLICATION NUMBER: 60/310,847  
; PRIOR FILING DATE: 2001-08-09  
; PRIOR APPLICATION NUMBER: 60/336,049



;; PRIOR FILING DATE: 2001-12-05  
;; PRIOR APPLICATION NUMBER: 60/338,692  
;; PRIOR FILING DATE: 2001-12-11  
;; PRIOR APPLICATION NUMBER: 10/171,468  
;; PRIOR FILING DATE: 2002-06-14  
;; NUMBER OF SEQ ID NOS: 1122  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 1084  
;; LENGTH: 664  
;; TYPE: PRT  
;; ORGANISM: Arabidopsis thaliana  
US-10-225-066A-1084

Query Match 71.4%; Score 35; DB 4; Length 664;  
Best Local Similarity 70.0%; Pred. No. 2.7e+02;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAVNHQKSA 10  
| | | | |  
Db 359 AKVEHERKSA 368

RESULT 7  
US-10-225-066A-1084  
; Sequence 1084, Application US/10225066A  
; Publication No. US20050160493A9  
; GENERAL INFORMATION:  
; APPLICANT: Mendel Biotechnology, Inc.  
; APPLICANT: RATCLIFFE, Oliver  
; APPLICANT: RIECHMANN, Jose Luis  
; APPLICANT: ADAM, Luc J  
; APPLICANT: DUBELL, Arnold T  
; APPLICANT: HEARD, Jacqueline E  
; APPLICANT: PILGRIM, Marsha L  
; APPLICANT: JIANG, Cai-Zhong  
; APPLICANT: REUBER, T. Lynne  
; APPLICANT: CREELMAN, Robert A  
; APPLICANT: PINEDA, Omaisra  
; APPLICANT: YU, Guo-Liang  
; APPLICANT: BROWN, Pierre E  
; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants  
; FILE REFERENCE: MBI0036-2 US  
; CURRENT APPLICATION NUMBER: US/10/225,066A  
; CURRENT FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: 09/837,444  
; PRIOR FILING DATE: 2001-04-18  
; PRIOR APPLICATION NUMBER: 60/310,847  
; PRIOR FILING DATE: 2001-08-09  
; PRIOR APPLICATION NUMBER: 60/336,049  
; PRIOR FILING DATE: 2001-12-05  
; PRIOR APPLICATION NUMBER: 60/338,692  
; PRIOR FILING DATE: 2001-12-11  
; PRIOR APPLICATION NUMBER: 10/171,468  
; PRIOR FILING DATE: 2002-06-14  
; NUMBER OF SEQ ID NOS: 1122  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1084  
; LENGTH: 664  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-10-225-066A-1084

Query Match 71.4%; Score 35; DB 5; Length 664;  
Best Local Similarity 70.0%; Pred. No. 2.7e+02;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAVNHQKSA 10  
| | | | |  
Db 359 AKVEHERKSA 368

RESULT 8  
US-10-369-493-19873  
; Sequence 19873, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493

;; PRIOR FILING DATE: 2001-12-05  
;; PRIOR APPLICATION NUMBER: 60/338,692  
;; PRIOR FILING DATE: 2001-12-11  
;; PRIOR APPLICATION NUMBER: 10/171,468  
;; PRIOR FILING DATE: 2002-06-14  
;; NUMBER OF SEQ ID NOS: 1122  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 1084  
;; LENGTH: 664  
;; TYPE: PRT  
;; ORGANISM: Arabidopsis thaliana  
US-10-225-066A-1084

Query Match 71.4%; Score 35; DB 4; Length 664;  
Best Local Similarity 70.0%; Pred. No. 2.7e+02;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAVNHQKSA 10  
| | | | |  
Db 359 AKVEHERKSA 368

RESULT 6  
US-10-374-780A-2530  
; Sequence 2530, Application US/10374780A  
; Publication No. US20040019927A1  
; GENERAL INFORMATION:  
; APPLICANT: Sherman, Bradley K  
; APPLICANT: Riechmann, Jose Luis  
; APPLICANT: Jiang, Cai-Zhong  
; APPLICANT: HEARD, Jacqueline E  
; APPLICANT: Haake, Volker  
; APPLICANT: Creelman, Robert A  
; APPLICANT: Ratcliffe, Oliver  
; APPLICANT: Adam, Luc J  
; APPLICANT: Reuber, T. Lynne  
; APPLICANT: Keddle, James  
; APPLICANT: BROWN, Pierre E  
; APPLICANT: Pilgrim, Marsha L  
; APPLICANT: Dubell III, Arnold T  
; APPLICANT: Pineda, Omaisra  
; APPLICANT: Yu, Guo-Liang  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS  
; FILE REFERENCE: MBI-0047 CIP  
; CURRENT APPLICATION NUMBER: US/10/374,780A  
; CURRENT FILING DATE: 2003-02-25  
; PRIOR APPLICATION NUMBER: 09/837,944  
; PRIOR FILING DATE: 2001-04-18  
; PRIOR APPLICATION NUMBER: 60/310,847  
; PRIOR FILING DATE: 2001-08-09  
; PRIOR APPLICATION NUMBER: 09/934,455  
; PRIOR FILING DATE: 2001-08-22  
; PRIOR APPLICATION NUMBER: 60/336,049  
; PRIOR FILING DATE: 2001-11-19  
; PRIOR APPLICATION NUMBER: 60/338,692  
; PRIOR FILING DATE: 2001-12-11  
; PRIOR APPLICATION NUMBER: 10/171,468  
; PRIOR FILING DATE: 2002-06-14  
; PRIOR APPLICATION NUMBER: 10/225,066  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: 10/225,067  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: 10/225,068  
; PRIOR FILING DATE: 2002-08-09  
; NUMBER OF SEQ ID NOS: 2906  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2530  
; LENGTH: 664  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; OTHER INFORMATION: G1079  
US-10-374-780A-2530

```
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 19873
; LENGTH: 284
; TYPE: PRT
; ORGANISM: No. US20030233675Altoc punctiforme
US-10-369-493-19873
```

```
Query Match          69.4%; Score 34; DB 4; Length 284;
Best Local Similarity 81.8%; Pred. No. 1.8e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 2; Gaps 1;
```

```
QY      1 AAVNH--QRKS 9
         ||||| |||||
Db       126 AAVNHFKQRKS 136
```

## RESULT 9

```
US-10-450-763-38089
; Sequence 38089, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: US/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 38089
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
; NAME/KEY: DOMAIN
; LOCATION: (143)..(289)
; OTHER INFORMATION: Reverse transcriptase domain identified by Pfam, accession
; OTHER INFORMATION: name rvt, E-value=5.8e-05, Pfam score of 0.9
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (143)..(289)
; OTHER INFORMATION: Reverse transcriptase domain identified by Pfam, accession
; OTHER INFORMATION: name rvt, E-value=9.135e-09, raw score of 11.58
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (240)..(252)
; OTHER INFORMATION: RNase H domain identified by eMATRIX, accession number
; OTHER INFORMATION: PF00075C, p-value=9.135e-09, raw score of 11.58
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (143)..(289)
; OTHER INFORMATION: Reverse transcriptase domain identified by Pfam, accession
US-10-450-763-38089
```

```
Query Match          69.4%; Score 34; DB 5; Length 291;
Best Local Similarity 75.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 AAVNHQRK 8
         ||:||||:
Db       66 AALNHQRQ 73
```

## RESULT 10

```
US-10-450-763-44774
; Sequence 44774, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
```

```
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: US/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 44774
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (228)..(244)
; OTHER INFORMATION: AUXIN BINDING PROTEIN SIGNATURE domain identified by eMATRIX,
; OTHER INFORMATION: accession number PR00655E, p-value=7.222e-09, raw score of 8.06
US-10-450-763-44774
```

```
Query Match          69.4%; Score 34; DB 5; Length 300;
Best Local Similarity 75.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 AAVNHQRK 8
         ||:||||:
Db       80 AALNHQRQ 87
```

## RESULT 11

```
US-10-450-763-35911
; Sequence 35911, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: US/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 35911
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (6)..(36)
; OTHER INFORMATION: kw TRANSCRIPTASE REVERSE II ORF2 domain identified by
; OTHER INFORMATION: eMATRIX, accession number DM01354R, p-values=8.948e-17, raw score
; OTHER INFORMATION: 8.50
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (273)..(434)
; OTHER INFORMATION: Reverse transcriptase domain identified by Pfam, accession
; OTHER INFORMATION: name rvt, E-value=4.9e-06, Pfam score of 16.1
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(445)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-35911
```

```
Query Match          69.4%; Score 34; DB 5; Length 445;
Best Local Similarity 75.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 AAVNHQRK 8
         ||:||||:
Db       196 AALNHQRQ 203
```

```
RESULT 12
US-10-450-763-44530
; Sequence 44530, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 44530
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (6)..(36)
; OTHER INFORMATION: kw TRANSCRIPTASE REVERSE II ORF2 domain identified by
; OTHER INFORMATION: eMATRIX, accession number DM01354R, p-value=8.948e-17, raw score
; OTHER INFORMATION: 8.50
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(453)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-44530

Query Match 69.4%; Score 34; DB 5; Length 453;
Best Local Similarity 75.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAVNHQK 8
Db 196 AALNHQK 203

RESULT 13
US-10-450-763-36492
; Sequence 36492, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 36492
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (6)..(36)
; OTHER INFORMATION: kw TRANSCRIPTASE REVERSE II ORF2 domain identified by
; OTHER INFORMATION: eMATRIX, accession number DM01354R, p-value=8.948e-17, raw score
; OTHER INFORMATION: 8.50
US-10-450-763-36492

Query Match 69.4%; Score 34; DB 5; Length 454;
Best Local Similarity 75.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAVNHQK 8
Db 196 AALNHQK 203

RESULT 14
US-10-425-115-201509
; Sequence 201509, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 201509
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_11535C.1.pep
US-10-425-115-201509

Query Match 67.3%; Score 33; DB 4; Length 57;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 NHQKRS 9
Db 46 NHQKRS 51

RESULT 15
US-10-425-115-253427
; Sequence 253427, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 253427
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_162707C.1.pep
US-10-425-115-253427

Query Match 67.3%; Score 33; DB 4; Length 108;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 NHQKSA 10
Db 60 NHQKSA 66
```

US-10-450-763-36492

Query Match 69.4%; Score 34; DB 5; Length 454;  
Best Local Similarity 75.0%; Pred. No. 2.8e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAVNHQK 8  
Db 196 AALNHQK 203

RESULT 14

US-10-425-115-201509  
; Sequence 201509, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 201509  
; LENGTH: 57  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_11535C.1.pep  
US-10-425-115-201509

Query Match 67.3%; Score 33; DB 4; Length 57;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 NHQKRS 9  
Db 46 NHQKRS 51

RESULT 15

US-10-425-115-253427  
; Sequence 253427, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 253427  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_162707C.1.pep  
US-10-425-115-253427

Query Match 67.3%; Score 33; DB 4; Length 108;  
Best Local Similarity 85.7%; Pred. No. 1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 NHQKSA 10  
Db 60 NHQKSA 66

Search completed: May 3, 2006, 19:17:52  
Job time : 5.50823 secs



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2006, 19:12:24 ; Search time 0.576132 Seconds  
(without alignments)

803.371 Million cell updates/sec

Title: US-10-525-567-1

Perfect score: 49

Sequence: 1 AAVNHQKSA 10

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 235405 seqs, 46284737 residues

Total number of hits satisfying chosen parameters: 235405

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:

- 1: /SID55/ptodata/1/pubpaa/US08 NEW PUB.pepl.\*
- 2: /SID55/ptodata/1/pubpaa/US06 NEW PUB.pep.\*
- 3: /SID55/ptodata/1/pubpaa/US07 NEW PUB.pep.\*
- 4: /SID55/ptodata/1/pubpaa/US08 NEW PUB.pep.\*
- 5: /SID55/ptodata/1/pubpaa/PCT NEW PUB.pep.\*
- 6: /SID55/ptodata/1/pubpaa/US09 NEW PUB.pep.\*
- 7: /SID55/ptodata/1/pubpaa/US09 NEW PUB.pep.\*
- 8: /SID55/ptodata/1/pubpaa/US10 NEW PUB.pep.\*
- 9: /SID55/ptodata/1/pubpaa/US10 NEW PUB.pep.\*
- 10: /SID55/ptodata/1/pubpaa/US11 NEW PUB.pep.\*
- 11: /SID55/ptodata/1/pubpaa/US11 NEW PUB.pep.\*
- 12: /SID55/ptodata/1/pubpaa/US60 NEW PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	67.3	284	11 US-11-079-463-5266	Sequence 1106, Ap
2	33	67.3	809	9 US-10-467-657-1106	Sequence 1106, Ap
3	33	67.3	810	11 US-11-052-554A-245	Sequence 245, App
4	32	65.3	184	11 US-11-087-099-5689	Sequence 5689, Ap
5	32	65.3	297	11 US-11-096-568A-18592	Sequence 18592, A
6	32	65.3	432	11 US-11-087-099-8419	Sequence 8419, Ap
7	32	65.3	432	11 US-11-188-298-7797	Sequence 7797, Ap
8	32	65.3	599	11 US-11-087-099-1241	Sequence 1241, Ap
9	31	63.3	38	9 US-10-986-501-354	Sequence 354, App
10	31	63.3	265	11 US-11-098-686-11404	Sequence 11404, A
11	30	61.2	133	9 US-10-467-657-5098	Sequence 5098, Ap
12	30	61.2	143	9 US-10-467-657-9090	Sequence 9090, Ap
13	30	61.2	237	9 US-10-714-887-254	Sequence 254, App
14	30	61.2	267	11 US-11-188-298-3220	Sequence 3220, Ap
15	30	61.2	278	11 US-11-098-686-11365	Sequence 11365, A
16	30	61.2	341	9 US-10-515-481-12	Sequence 12, Appl
17	30	61.2	341	11 US-11-087-099-1836	Sequence 1836, Ap
18	30	61.2	341	11 US-11-045-004-871	Sequence 871, App
19	30	61.2	1159	9 US-10-055-877-139	Sequence 139, App
20	30	61.2	1231	11 US-11-096-568A-28514	Sequence 28514, A
21	30	61.2	1253	11 US-11-096-568A-28513	Sequence 28513, A

#### ALIGNMENTS

##### RESULT 1

US-11-079-463-5266

Sequence 5266, Application US/11079463

Publication No. US20060073161A1

GENERAL INFORMATION:

APPLICANT: Gary L. Breton

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FR

FILE REFERENCE: PATH00-03DIV2

CURRENT APPLICATION NUMBER: US/11/079,463

CURRENT FILING DATE: 2005-03-14

PRIOR APPLICATION NUMBER: US 60/128,705

PRIOR FILING DATE: 1999-04-09

PRIOR APPLICATION NUMBER: US 09/540,209

PRIOR FILING DATE: 2000-04-04

NUMBER OF SEQ ID NOS: 10444

SEQ ID NO 5266

LENGTH: 284

TYPE: PRT

ORGANISM: B.fragilis

US-11-079-463-5266

Query Match 67.3%; Score 33; DB 11; Length 284;

Best Local Similarity 87.5%; Pred. No. 37;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAVNHQK 8

Db 135 AAVNEQK 142

##### RESULT 2

US-10-467-657-1106

Sequence 1106, Application US/10467657

Publication No. US20050260581A1

GENERAL INFORMATION:

APPLICANT: CHIRON SpA

APPLICANT: FONTANA Maria Rita

APPLICANT: PIZZA Mariagrazia

APPLICANT: MASIGNANI Vega

APPLICANT: MONACI Elisabetta

TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS

FILE REFERENCE:

CURRENT APPLICATION NUMBER: US/10/467,657

Sequence 28512, A  
Sequence 8532, Ap  
Sequence 4, Appl  
Sequence 17240, A  
Sequence 393, App  
Sequence 17239, A  
Sequence 25, Appl  
Sequence 565, App  
Sequence 11897, A  
Sequence 10824, A  
Sequence 6115, Ap  
Sequence 17592, A  
Sequence 10573, A  
Sequence 4026, Ap  
Sequence 2, Appl  
Sequence 13, Appl  
Sequence 69, Appl  
Sequence 10572, A  
Sequence 9031, Ap  
Sequence 8113, Ap  
Sequence 19571, A  
Sequence 19589, A  
Sequence 2947, Ap  
Sequence 638, App

22 61.2 1401 11 US-11-096-568A-28512  
23 59.2 76 9 US-10-467-657-8532  
24 59.2 92 11 US-11-172-274-4  
25 59.2 236 11 US-11-096-568A-17240  
26 59.2 252 11 US-11-045-004-393  
27 59.2 254 11 US-11-096-568A-17239  
28 59.2 261 9 US-10-537-897-25  
29 59.2 261 11 US-11-172-740-565  
30 59.2 261 11 US-11-188-298-11897  
31 59.2 267 11 US-11-096-568A-10824  
32 59.2 302 11 US-11-188-298-6115  
33 59.2 309 11 US-11-188-298-17592  
34 59.2 331 11 US-11-096-568A-10573  
35 59.2 332 9 US-10-467-657-4026  
36 59.2 337 9 US-10-875-716-2  
37 59.2 337 11 US-11-174-816-13  
38 59.2 337 11 US-11-174-819-69  
39 59.2 338 11 US-11-096-568A-10572  
40 59.2 350 11 US-11-079-463-9031  
41 59.2 351 11 US-11-188-298-8113  
42 59.2 381 11 US-11-096-568A-10571  
43 59.2 416 11 US-11-188-298-19589  
44 59.2 441 8 US-10-511-937-2947  
45 59.2 441 9 US-10-995-561-638

```
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 1106
; LENGTH: 809
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1106

Query Match      67.3%; Score 33; DB 9; Length 809;
Best Local Similarity 77.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AVNHQKSA 10
   |||:|:|
Db 160 AVNYQTKSA 168

RESULT 3
US-11-052-554A-245
; Sequence 245, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 245
; LENGTH: 810
; TYPE: PRT
; ORGANISM: Neisseria meningitidis Z2491
US-11-052-554A-245

Query Match      67.3%; Score 33; DB 11; Length 810;
Best Local Similarity 77.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AVNHQKSA 10
   |||:|:|
Db 160 AVNYQTKSA 168

RESULT 4
US-11-087-099-5689
; Sequence 5689, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 5689
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(184)
; OTHER INFORMATION: unsure at all Xaa locations
US-11-087-099-5689
```

```
Query Match      65.3%; Score 32; DB 11; Length 184;
Best Local Similarity 75.0%; Pred. No. 37;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VNHQKSA 10
   |||:|:|
Db 166 VNHQKIA 173

RESULT 5
US-11-096-568A-18592
; Sequence 18592, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 18592
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(297)
; OTHER INFORMATION: Ceres Seq. ID no. 12365991
US-11-096-568A-18592

Query Match      65.3%; Score 32; DB 11; Length 297;
Best Local Similarity 66.7%; Pred. No. 62;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAVNHQKRS 9
   |||:|:|
Db 230 AAAEHQKKA 238

RESULT 6
US-11-087-099-8419
; Sequence 8419, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 8419
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(432)
; OTHER INFORMATION: unsure at all Xaa locations
US-11-087-099-8419

Query Match      65.3%; Score 32; DB 11; Length 432;
Best Local Similarity 75.0%; Pred. No. 93;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AVNHQKRS 9
   |||:|:|
Db 213 AVDHQKRS 220

RESULT 7
US-11-188-298-7797
; Sequence 7797, Application US/11188298
```

; Publication No. US20060075522A1  
; GENERAL INFORMATION:  
; APPLICANT: Abad, Mark S. et al.  
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT  
; FILE REFERENCE: 38-21(53452)B  
; CURRENT APPLICATION NUMBER: US/11/188,298  
; CURRENT FILING DATE: 2005-07-22  
; PRIOR APPLICATION NUMBER: 60/592,978  
; PRIOR FILING DATE: 2004-07-31  
; NUMBER OF SEQ ID NOS: 22569  
; SEQ ID NO 7797  
; LENGTH: 432  
; TYPE: PRT  
; ORGANISM: Myxococcus xanthus  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(432)  
; OTHER INFORMATION: unsure at all Xaa locations  
US-11-188-298-7797

Query Match 65.3%; Score 32; DB 11; Length 432;  
Best Local Similarity 75.0%; Pred. No. 93;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AVNHQKS 9  
|:|:|  
Db 213 AVDHQKS 220

RESULT 8  
US-11-087-099-1241  
; Sequence 1241, Application US/11087099  
; Publication No. US20060041961A1  
; GENERAL INFORMATION:  
; APPLICANT: Abad, Mark S. et al.  
; TITLE OF INVENTION: Genes and Uses for Plant Improvement  
; FILE REFERENCE: 38-21(53450)B EP  
; CURRENT APPLICATION NUMBER: US/11/087,099  
; CURRENT FILING DATE: 2005-03-22  
; NUMBER OF SEQ ID NOS: 12464  
; SEQ ID NO 1241  
; LENGTH: 599  
; TYPE: PRT  
; ORGANISM: Coxsiella burnetii RSA 493  
US-11-087-099-1241

Query Match 65.3%; Score 32; DB 11; Length 599;  
Best Local Similarity 70.0%; Pred. No. 1.3e+02;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAVNHQKSA 10  
|:|:|  
Db 534 AAVNHQKSA 543

RESULT 9  
US-10-986-501-354  
; Sequence 354, Application US/10986501  
; Publication No. US2005024845A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 90 Human Secreted Proteins  
; FILE REFERENCE: P2013P2C1  
; CURRENT APPLICATION NUMBER: US/10/986,501  
; CURRENT FILING DATE: 2004-11-12  
; PRIOR APPLICATION NUMBER: US/10/621,363  
; PRIOR FILING DATE: 2003-07-18  
; PRIOR APPLICATION NUMBER: 09/969,730  
; PRIOR FILING DATE: 2001-10-06  
; PRIOR APPLICATION NUMBER: 09/774,639  
; PRIOR FILING DATE: 2001-02-01  
; PRIOR APPLICATION NUMBER: 60/238,291  
; PRIOR FILING DATE: 2000-10-06

; PRIOR APPLICATION NUMBER: 09/244,112  
; PRIOR FILING DATE: 1999-02-04  
; PRIOR APPLICATION NUMBER: PCT/US98/16235  
; PRIOR FILING DATE: 1998-08-04  
; PRIOR APPLICATION NUMBER: 60/056,371  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 60/056,732  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 60/056,366  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 60/056,364  
; PRIOR FILING DATE: 1997-08-19  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 373  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 354  
; LENGTH: 38  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-986-501-354

Query Match 63.3%; Score 31; DB 9; Length 38;  
Best Local Similarity 62.5%; Pred. No. 11;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VNHQKSA 10  
|:|:|  
Db 1 IRHERKSA 8

RESULT 10  
US-11-098-686-11404  
; Sequence 11404, Application US/11098686  
; Publication No. US20060024696A1  
; GENERAL INFORMATION:  
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.  
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES  
; FILE REFERENCE: 09531-128001  
; CURRENT APPLICATION NUMBER: US/11/098,686  
; CURRENT FILING DATE: 2005-04-04  
; PRIOR APPLICATION NUMBER: PCT/US03/31318  
; PRIOR FILING DATE: 2003-10-01  
; PRIOR APPLICATION NUMBER: US 60/416,395  
; PRIOR FILING DATE: 2002-10-04  
; NUMBER OF SEQ ID NOS: 11433  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11404  
; LENGTH: 265  
; TYPE: PRT  
; ORGANISM: Lawsonia intracellularis  
US-11-098-686-11404

Query Match 63.3%; Score 31; DB 11; Length 265;  
Best Local Similarity 66.7%; Pred. No. 86;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AVNHQKSA 10  
|:|:|  
Db 103 AMNQKAA 111

RESULT 11  
US-10-467-657-5098  
; Sequence 5098, Application US/10467657  
; Publication No. US20050260581A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON SpA  
; APPLICANT: FONTANA Maria Rita  
; APPLICANT: PIZZA Mariagrazia  
; APPLICANT: MASIGNANI Vega  
; APPLICANT: MONACI Elisabetta  
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS

```
; FILE REFERENCE: MBI0058-CIP
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 5098
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-5098

Query Match          61.2%; Score 30; DB 9; Length 133;
Best Local Similarity 62.5%; Pred. No. 65;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VNHQKSA 10
Db 72 LNQHRRPA 79

RESULT 12
US-10-467-657-9090
; Sequence 9090, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 9090
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-9090

Query Match          61.2%; Score 30; DB 9; Length 143;
Best Local Similarity 71.4%; Pred. No. 70;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 NHQKSA 10
Db 45 NHQRQTA 51

RESULT 13
US-10-714-887-254
; Sequence 254, Application US/10714887
; Publication No. US20060015972A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: HEARD, Jacqueline
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: CREELMAN, Robert
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: CANALES, Roger
; APPLICANT: REPETTI, Peter
; APPLICANT: KUMIMOTO, Roderick W
; APPLICANT: GUTTERSON, Neal
; APPLICANT: REUBER, T. Lynne
; APPLICANT: PINEDA, Omalra
; APPLICANT: SHERMAN, Bradley K
; TITLE OF INVENTION: PLANT TRANSCRIPTIONAL REGULATORS OF DROUGHT STRESS

; FILE REFERENCE: MBI0058-CIP
; CURRENT APPLICATION NUMBER: US/10/714,887
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: 10/412,699
; PRIOR FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/135,134
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/125,814
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 430
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 254
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G3083 polypeptide; reference sequence
US-10-714-887-254

Query Match          61.2%; Score 30; DB 9; Length 237;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAVNHQKSA 10
Db 30 SAQNHHRNSA 39

RESULT 14
US-11-188-298-3220
; Sequence 3220, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 3220
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Oryza sativa (japonica cultivar-group)
US-11-188-298-3220

Query Match          61.2%; Score 30; DB 11; Length 267;
Best Local Similarity 60.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAVNHQKSA 10
Db 20 AAVHHNRSA 29

RESULT 15
US-11-098-686-11365
```



Thu May 4 17:23:15 2006

```
; Sequence 11365, Application US/11098686
; Publication No. US20060024698A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11365
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-11365
```

```
Query Match          61.2%; Score 30; DB 11; Length 278;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy      1 AAVNHQKSA 10
      ||:||:|
Db      246 AAINHMDRTA 255
```

```
Search completed: May 3, 2006, 19:18:54
Job time : 1.57613 secs
```

This Page Blank (usp10)

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2006, 18:53:48 ; Search time 10.6224 Seconds  
(without alignments)  
1034.083 Million cell updates/sec

Title: US-10-525-567-2

Perfect score: 138

Sequence: 1 NLRGSLNKSPLILHEWKFFDYDF 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_21.\*

- 1: Geneseqp1980s.\*
- 2: Geneseqp1990s.\*
- 3: Geneseqp2000s.\*
- 4: Geneseqp2001s.\*
- 5: Geneseqp2002s.\*
- 6: Geneseqp2003as.\*
- 7: Geneseqp2003bs.\*
- 8: Geneseqp2004s.\*
- 9: Geneseqp2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	138	100.0	25	8 ADL27605	Adl27605 Honey bee
2	138	100.0	413	8 ADL27607	Adl27607 Honey bee
3	138	100.0	414	8 ADL27617	Adl27617 Honey bee
4	138	100.0	432	4 AAG78710	Aag78710 Oxygen pr
5	138	100.0	432	4 AAG78815	Aag78815 Bee prote
6	138	100.0	432	8 ADL27619	Adl27619 Honey bee
7	101	73.2	25	4 AAG78709	Aag78709 Oxygen pr
8	101	73.2	25	4 AAG78814	Aag78814 Bee pepti
9	101	73.2	25	5 AAO20913	Aao20913 25-mer Ap
10	101	73.2	25	5 AAO14842	Aao14842 Honeybee
11	101	73.2	25	7 ADC52131	Adc52131 Apis mell
12	101	73.2	25	8 ADL06420	Adl06420 Protein-k
13	101	73.2	25	8 ADN61432	Adn61432 Honey bee
14	55.5	40.2	768	2 AA94918	Aaw94918 Rat phero
15	55	39.9	524	8 ADL27606	Adl27606 Honey bee
16	55	39.9	544	8 ADL27618	Adl27618 Honey bee
17	55	39.9	544	8 ADL27616	Adl27616 Honey bee
18	52	37.7	9	2 AAR28285	Aar28285 RJP3 - AC
19	52	37.7	42	4 ABB03089	Abb03089 Human mus
20	52	37.7	42	6 ABU12383	Abu12383 Novel hum
21	52	37.7	42	8 ADJ28409	Adj28409 Human mus
22	51	37.0	281	6 ABM70854	Abm70854 Staphyloc
23	50	36.2	116	8 AOG65186	Aog65186 Novel hum
24	49.5	35.9	255	6 ABU16743	Abu16743 Protein e

25	49.5	35.9	270	8 ADP29330	Adp29330 Human sec
26	49.5	35.9	411	6 ADA33752	Ada33752 Acinetoba
27	49.5	35.9	610	6 ADA36517	Ada36517 Acinetoba
28	49.5	35.9	1029	6 ABR53179	Abr53179 Protein s
29	49.5	35.9	1029	7 ADK63210	Adk63210 Disease t
30	49	35.5	142	7 ADG30689	Adg30689 Xanthomon
31	49	35.5	768	5 ABB49385	Abb49385 Listeria
32	49	35.5	768	6 ABU32627	Abu32627 Protein e
33	48	34.8	400	8 ADN47954	Adn47954 Thermococ
34	48	34.8	445	8 ADX66511	Adx66511 Plant ful
35	48	34.8	904	5 ABP74080	Abp74080 Candida a
36	47.5	34.4	949	5 ABP35671	Abp35671 Fungal ZB
37	47.5	34.4	204	3 AAB08470	Aab08470 Amino aci
38	47	34.1	85	7 ADD25195	Add25195 Fertility
39	47	34.1	85	8 ADN61210	Adn61210 Radish nu
40	47	34.1	109	4 AAO00965	Aao00965 Human pol
41	47	34.1	242	8 ADX90871	Adx90871 Plant ful
42	47	34.1	242	8 ADX75685	Adx75685 Plant ful
43	47	34.1	380	4 AAU09147	Aau09147 Vasodilat
44	47	34.1	380	7 ADG46667	Adg46667 Human vas
45	47	34.1	380	8 ADJ14192	Adj14192 Human (va

## ALIGNMENTS

### RESULT 1

ADL27605

ID ADL27605 standard; peptide; 25 AA.

XX ADL27605;

XX 03-JUN-2004 (first entry)

XX Honey bee RJP55 N-terminal peptide, SEQ ID 2.

XX Antiallergic; Antiasthmatic; Antiinflammatory; Dermatological; honey bee;  
royal jelly protein; allergy; pollinosis; atopic dermatitis;  
contact hypersensitivity; bronchial asthma; allergic rhinitis; RJP55.

OS Apis mellifera.

XX WO2004019971-A1.

XX 11-MAR-2004.

XX 26-AUG-2003; 2003WO-JP010795.

XX 29-AUG-2002; 2002JP-00252087.

XX 30-JAN-2003; 2003JP-00022776.

XX (HAYA/) HAYASHIBARA K.

XX Okamoto I, Arai N, Kohno K, Kurimoto M, Sano O;

XX WPI; 2004-248191/23.

XX Antiallergic agents for relieving symptoms accompanying e.g. pollinosis,  
atopic dermatitis, contact hypersensitivity, bronchial asthma and  
allergic rhinitis, comprise protein originated from royal jelly.

XX Claim 2; SEQ ID NO 2; 78pp; Japanese.

XX The present invention relates to novel antiallergic agents, which  
comprise as an active ingredient, peptides ADL27604 and ADL27605, which  
are N-terminal peptides of honey bee royal jelly proteins (RJP, ADL27606  
and ADL27607). The agents can be used to relieve symptoms accompanying an  
allergic disease e.g. pollinosis, atopic dermatitis, contact  
hypersensitivity, bronchial asthma and allergic rhinitis, which are  
applicable in foods, cosmetics and drugs.

XX Sequence 25 AA;

```
Query Match      100.0%; Score 138; DB 8; Length 25;
Best Local Similarity 100.0%; Pred. No. 4e-14;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NILRGSLNKSLLPILHEWKFFDYDF 25
   |||||
Db 1 NILRGSLNKSLLPILHEWKFFDYDF 25

RESULT 2
ADL27607
ID ADL27607 standard; protein; 413 AA.
XX AC ADL27607;
XX AC
DT 03-JUN-2004 (first entry)
DE Honey bee RJP55 mature protein, SEQ ID 4.
XX ADL27607;
XX ADL27607;
KW Antiallergic; Antiasthmatic; Antiinflammatory; Dermatological; honey bee;
KW royal jelly protein; allergy; pollinosis; atopic dermatitis;
KW contact hypersensitivity; bronchial asthma; allergic rhinitis; RJP55.
XX OS Apis mellifera.
XX WO2004019971-A1.
XX PN
XX PD 11-MAR-2004.
XX PF 26-AUG-2003; 2003WO-JP010795.
XX PR 29-AUG-2002; 2002JP-00252087.
XX PR 30-JAN-2003; 2003JP-00022776.
XX PR (HAYA/) HAYASHIBARA K.
XX PA
XX PI Okamoto I, Arai N, Kohno K, Kurimoto M, Sano O;
XX PI
XX DR WPI: 2004-248191/23.
XX DR N-PSDB; ADL27609.
XX DR
XX PT Antiallergic agents for relieving symptoms accompanying e.g. pollinosis,
XX PT atopic dermatitis, contact hypersensitivity, bronchial asthma and
XX PT allergic rhinitis, comprise protein originated from royal jelly.
XX PS Disclosure; Page 62-65; 78pp; Japanese.
XX PS
XX CC The present invention relates to novel antiallergic agents, which
XX CC comprise as an active ingredient, peptides ADL27604 and ADL27605, which
XX CC are N-terminal peptides of honey bee royal jelly proteins (RJP, ADL27606
XX CC and ADL27607). The agents can be used to relieve symptoms accompanying an
XX CC allergic disease e.g. pollinosis, atopic dermatitis, contact
XX CC hypersensitivity, bronchial asthma and allergic rhinitis, which are
XX CC applicable in foods, cosmetics and drugs. The present sequence is a honey
XX CC bee RJP.
XX SQ Sequence 413 AA;

Query Match      100.0%; Score 138; DB 8; Length 414;
Best Local Similarity 100.0%; Pred. No. 9e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NILRGSLNKSLLPILHEWKFFDYDF 25
   |||||
Db 2 NILRGSLNKSLLPILHEWKFFDYDF 26

RESULT 4
AAG78710
ID AAG78710 standard; protein; 432 AA.
XX AC
XX AC AAG78710;
XX AC
DT 11-DEC-2001 (first entry)
XX DT
XX DE Oxygen promoter related protein.
XX DE
XX KW Oxygen promoter; bee; royal jelly; cardiant; vulnerary;
XX KW cardiopulmonary function.
XX KW
XX OS Apis mellifera.
XX OS
XX PN JP2001172195-A.
XX PN
XX PD 26-JUN-2001.
XX PD
XX PF 15-DEC-1999; 99JP-00356005.
XX PF
XX PR 15-DEC-1999; 99JP-00356005.
```

```

XX PA (POKK ) POLA CHEM IND INC.
XX WPI; 2001-592558/67.
XX N-PSDB; AAI64996.
XX Agent for promoting oxygen in cells to improve cardiopulmonary function,
XX comprises protein in royal jelly which comprises oxygen enhancing effect.
XX Disclosure; Page 9-10; 10pp; Japanese.
XX The present invention relates to an agent for promoting oxygen,
XX comprising a protein in royal jelly which has an oxygen enhancing effect.
XX This can be used to promote oxygen in cells, causing an improvement in
XX pulmonary function. The present sequence is a protein which can be used
XX in the agent of the invention
XX SQ Sequence 432 AA;
Query Match 100.0%; Score 138; DB 4; Length 432;
Best Local Similarity 100.0%; Pred. No. 9.4e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NILRGESLNKSLPILHEWKFFDYDF 25
Db 20 NILRGESLNKSLPILHEWKFFDYDF 44

RESULT 6
ADL27619
ID ADL27619 standard; protein; 432 AA.
XX AC ADL27619;
XX DT 03-JUN-2004 (first entry)
XX DE Honey bee MRJP1.
XX KW Antiallergic; Antiasthmatic; Antiinflammatory; Dermatological; honey bee;
XX KW royal jelly protein; allergy; pollinosis; atopic dermatitis;
XX KW contact hypersensitivity; bronchial asthma; allergic rhinitis; MRJP1.
XX OS Apis mellifera.
XX FH Key Location/Qualifiers
XX FT Peptide 1..19
XX FT Protein /label= Signal_peptide
XX FT FT 20..432
XX FT FT /label= Mature_protein
XX PN WO2004019971-A1.
XX PD 11-MAR-2004.
XX PF 26-AUG-2003; 2003WO-JP010795.
XX PR 29-AUG-2002; 2002JP-00252087.
XX PR 30-JAN-2003; 2003JP-00022776.
XX PA (HAYA/) HAYASHIBARA K.
XX PI Okamoto I, Arai N, Kohno K, Kurimoto M, Sano O;
XX WPI; 2004-248191/23.
XX DR N-PSDB; ADL27611.
XX Antiallergic agents for relieving symptoms accompanying e.g. pollinosis,
XX atopic dermatitis, contact hypersensitivity, bronchial asthma and
XX allergic rhinitis, comprise protein originated from royal jelly.
XX Disclosure; Page 70-72; 78pp; Japanese.
XX The present invention relates to novel antiallergic agents, which
XX comprise as an active ingredient, peptides ADL27604 and ADL27605, which
XX are N-terminal peptides of honey bee royal jelly proteins (RJP, ADL27606
XX and ADL27607). The agents can be used to relieve symptoms accompanying an
XX allergic disease e.g. pollinosis, atopic dermatitis, contact
XX hypersensitivity, bronchial asthma and allergic rhinitis, which are
XX applicable in foods, cosmetics and drugs. The present sequence is a honey
XX bee major RJP.
XX SQ Sequence 432 AA;
Query Match 100.0%; Score 138; DB 8; Length 432;
Best Local Similarity 100.0%; Pred. No. 9.4e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX PA (POKK ) POLA CHEM IND INC.
XX WPI; 2001-592558/67.
XX N-PSDB; AAI64996.
XX Agent for promoting oxygen in cells to improve cardiopulmonary function,
XX comprises protein in royal jelly which comprises oxygen enhancing effect.
XX Disclosure; Page 9-10; 10pp; Japanese.
XX The present invention relates to an agent for promoting oxygen,
XX comprising a protein in royal jelly which has an oxygen enhancing effect.
XX This can be used to promote oxygen in cells, causing an improvement in
XX pulmonary function. The present sequence is a protein which can be used
XX in the agent of the invention
XX SQ Sequence 432 AA;
Query Match 100.0%; Score 138; DB 4; Length 432;
Best Local Similarity 100.0%; Pred. No. 9.4e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NILRGESLNKSLPILHEWKFFDYDF 25
Db 20 NILRGESLNKSLPILHEWKFFDYDF 44

RESULT 5
AAG78815
ID AAG78815 standard; protein; 432 AA.
XX AC AAG78815;
XX DT 27-NOV-2001 (first entry)
XX DE Bee protein.
XX KW Bee; lactic acid accumulation inhibitor; royal jelly; muscular fatigue.
XX OS Apis mellifera.
XX FH Key Location/Qualifiers
XX FT Peptide 1..19
XX FT Protein /label= Signal_peptide
XX FT Protein 20..432
XX FT Protein /label= Mature_bee_protein
XX FT Misc-difference 363
XX FT /label= unknown
XX FT /note= "Encoded by CTN, given as Xaa in specification"
XX JP2001172190-A.
XX 26-JUN-2001.
XX PF 15-DEC-1999; 99JP-00356006.
XX PR 15-DEC-1999; 99JP-00356006.
XX (POKK ) POLA CHEM IND INC.
XX WPI; 2001-592556/67.
XX N-PSDB; AAI65041.
XX Lactic acid accumulation inhibitor comprises protein contained in royal
XX jelly, useful for restoration, improvement and release of muscular
XX fatigue.
XX Disclosure; Page 9-10; 10pp; Japanese.
XX The present invention relates to a lactic acid accumulation inhibitor.
XX The inhibitor comprises a protein contained in royal jelly having lactic
XX acid accumulation inhibiting activity as the active component. The lactic

```

```
QY      1 NILRGESLNKSLPILHEWKFFDYDF 25
Db      20 NILRGESLNKSLPILHEWKFFDYDF 44

RESULT 7
AAG78709
ID      AAG78709 standard; peptide; 25 AA.
XX
AC      AAG78709;
XX
DT      11-DEC-2001 (first entry)
XX
DE      Oxygen promoter related peptide.
XX
KW      Oxygen promoter; bee; royal jelly; cardiant; vulnery;
KW      cardiopulmonary function.
XX
OS      Apis mellifera.
XX
PN      JP2001172195-A.
XX
PD      26-JUN-2001.
XX
PF      15-DEC-1999; 99JP-00356005.
XX
PR      15-DEC-1999; 99JP-00356005.
XX
PA      (POKK ) POLA CHEM IND INC.
XX
WPI; 2001-592558/67.
XX
The present invention relates to an agent for promoting oxygen,
comprising a protein in royal jelly which has an oxygen enhancing effect.
This can be used to promote oxygen in cells, causing an improvement in
pulmonary function. The present sequence is a peptide which can be used
in the agent of the invention
XX
Sequence 25 AA;
Query Match      73.2%; Score 101; DB 4; Length 25;
Best Local Similarity 87.0%; Pred. No. 2.4e-08;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 NILRGESLNKSLPILHEWKFFDY 23
Db      1 NILRGESLNKSLPILHEWKFFDY 23

RESULT 8
AAG78814
ID      AAG78814 standard; peptide; 25 AA.
XX
AC      AAG78814;
XX
DT      27-NOV-2001 (first entry)
XX
DE      Bee peptide.
XX
KW      Bee; lactic acid accumulation inhibitor; royal jelly; muscular fatigue.
XX
OS      Apis mellifera.
XX
Key      Location/Qualifiers
FT      Misc-difference 24
FT      /note= "Unspecified residue"
XX

QY      1 NILRGESLNKSLPILHEWKFFDY 23
Db      1 NILRGESLNKSLPILHEWKFFDY 23

RESULT 9
AAO20913
ID      AAO20913 standard; peptide; 25 AA.
XX
AC      AAO20913;
XX
DT      08-JUL-2002 (first entry)
XX
DE      25-mer Apis mellifera peptide.
XX
KW      Antiapoptotic; hepatotropic; apoptosis inhibitor; royal jelly.
XX
OS      Apis mellifera.
XX
Key      Location/Qualifiers
FT      Misc-difference 24
FT      /label= unknown
XX
PN      JP2002080391-A.
XX
PD      19-MAR-2002.
XX
PF      04-SEP-2000; 2000JP-00266421.
XX
PR      04-SEP-2000; 2000JP-00266421.
XX
PA      (POKK ) POLA CHEM IND INC.
XX
WPI; 2002-357223/39.
XX
An apoptosis inhibitor and a composition for hepatitis containing it.
XX
Claim 1; Page 5; 6pp; Japanese.
XX
The invention relates to an apoptosis inhibitor consisting of a protein
contained in royal jelly. It forms a single band in an unmodified
polyacrylamide gel electrophoresis of the protein in royal jelly. Its
```

CC molecular weight is 57 kDa when measured by sodium dodecyl sulfate (SDS)  
 CC -PAGE under reductive condition. It contains a 25 residue amino acid  
 CC sequence, given in the specification. The composition is used for  
 CC inhibiting apoptosis. This sequence represents the 25 residue amino acid  
 CC sequence of the invention  
 CC  
 XX  
 SQ Sequence 25 AA;

Query Match 73.2%; Score 101; DB 5; Length 25;  
 Best Local Similarity 87.0%; Pred. No. 2.4e-08;  
 Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NILRGESLNKSLPILHEWKFPDY 23  
 Db 1 NILRGESLLKKLPILHEWKFPDY 23

RESULT 10  
 AA014842  
 ID AA014842 standard; peptide; 25 AA.

XX  
 AC AA014842;

XX  
 DT 11-JUL-2002 (first entry)

XX Honeybee cell growth promoter-related peptide.

XX Honeybee; cell growth promoter; royal jelly; organ regeneration;  
 KW liver regeneration.

XX Apis mellifera.

Key Location/Qualifiers  
 FT Misc-difference 24  
 FT /label= Unknown

XX JP2002080392-A.

XX 19-MAR-2002.

XX 04-SEP-2000; 2000JP-00266422.

XX 04-SEP-2000; 2000JP-00266422.

XX (POKK ) POLA CHEM IND INC.

XX WPI; 2002-357224/39.

XX A cell growth promoter and a composition for organ regeneration  
 PT containing it.

XX Claim 1; Page 6; 7pp; Japanese.

XX The invention comprises a cell growth promoter protein present in royal  
 CC jelly. The invention also comprises a composition (containing the novel  
 CC cell growth promoter) for organ regeneration. The cell growth promoter is  
 CC useful in regenerating the liver. The present amino acid sequence  
 CC represents a honeybee cell growth promoter-related protein

XX Sequence 25 AA;

Query Match 73.2%; Score 101; DB 5; Length 25;  
 Best Local Similarity 87.0%; Pred. No. 2.4e-08;  
 Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NILRGESLNKSLPILHEWKFPDY 23

Db 1 NILRGESLLKKLPILHEWKFPDY 23

RESULT 11  
 ADC52131  
 ID ADC52131 standard; peptide; 25 AA.

XX ADC52131;  
 AC  
 XX 19-DEC-2003 (first entry)  
 DT  
 XX Apis mellifera peptide SEQ ID NO 1.  
 DE  
 XX Protein-kinase B; glycoprotein; royal jelly; liver;  
 KW congenital urea-cycle enzyme deficiency; hepatitis; alcoholic cirrhosis;  
 KW Apis mellifera; hepatotropic; antialcoholic; virucide; antiinflammatory;  
 KW Protein-kinase-stimulator-B; honeybee.  
 XX  
 OS Apis mellifera.  
 XX  
 XX Key Location/Qualifiers  
 FT Misc-difference 24  
 FT /note= "Xaa = unknown"

JP2003061648-A.

04-MAR-2003.

29-AUG-2001; 2001JP-00259187.

29-AUG-2001; 2001JP-00259187.

(POKK ) POLA CHEM IND INC.

WPI; 2003-580861/55.

Protein-kinase B activator for use in pharmaceutical or foodstuffs for  
 improving and recovering abnormal liver function and hepatitis, comprises  
 a glycoprotein obtained from royal jelly.

Disclosure; SEQ ID NO 1; 6pp; Japanese.

The invention relates to a protein-kinase B activator, consisting of a  
 glycoprotein obtained from royal jelly. The protein-kinase B activator is  
 used in a composition such as pharmaceutical or foodstuffs for improving  
 and recovering the abnormality of a liver function, congenital urea-cycle  
 enzyme deficiency, acute hepatitis, chronic-hepatitis, alcoholic fatty  
 liver, alcoholic cirrhosis, etc. The present sequence is that of an Apis  
 mellifera (honeybee) peptide disclosed as part of the invention.

Sequence 25 AA;

Query Match 73.2%; Score 101; DB 7; Length 25;

Best Local Similarity 87.0%; Pred. No. 2.4e-08;

Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NILRGESLNKSLPILHEWKFPDY 23

Db 1 NILRGESLLKKLPILHEWKFPDY 23

RESULT 12

ADL06420

ID ADL06420 standard; peptide; 25 AA.

XX  
 AC ADL06420;

DT 06-MAY-2004 (first entry)

XX Protein-kinase C activator associated Apis mellifera peptide.

XX Protein-kinase C activator; glycoprotein; royal jelly;  
 KW protein-kinase activation; foodstuff; myocardial infarction;  
 KW peroxide damage; cardiac; blood vessel; cardiacant.

XX Apis mellifera.

XX Key Location/Qualifiers  
 FT Misc-difference 24

```

FT XX /label= Unknown
PN XX JF2003313777-A.
PD XX 11-NOV-2003.
XX XX
XX XX 07-MAY-2002; 2002JP-00131184.
PF XX
XX XX 07-MAY-2002; 2002JP-00131184.
PR XX
XX XX (POKK ) POLA CHEM IND INC.
PA XX
XX XX WPI; 2004-085058/09.
XX XX
XX XX Protein-kinase C activator for use in composition such as foodstuffs to
PT PT prevent myocardial infarction, consists of glycoprotein having specified
PT PT molecular weight.
XX XX
PS Disclosure; SEQ ID NO 1; 8pp; Japanese.
XX XX
CC The present invention relates to a protein-kinase C activator consisting
CC of a glycoprotein which exists in royal jelly and forms a single band in
CC non-denaturing polyacrylamide gel electrophoresis. The glycoprotein has
CC molecular weight of 57 kilodalton. Also disclosed is a composition for
CC protein-kinase activation. The protein-kinase C activator is useful in
CC foodstuffs for preventing myocardial infarction. The protein-kinase C
CC activator provides protective effect from peroxide damage on cardiac or
CC blood vessels without any side effects. The present sequence represents
CC an Apis mellifera peptide.
XX XX
XX Sequence 25 AA;
SQ

Query Match 73.2%; Score 101; DB 8; Length 25;
Best Local Similarity 87.0%; Pred. No. 2.4e-08;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NILRGESLNKSLPILHEWKFFDY 23
Db 1 NILRGESLLKLPILHEWKFFDY 23

RESULT 13
ADN61432
ID ADN61432 standard; peptide; 25 AA.
XX
XX AC ADN61432;
XX XX
XX XX 01-JUL-2004 (first entry)
XX XX
XX XX Honey bee peptide of the serum-free culture medium SeqID 1.
XX XX
XX XX serum-free culture medium; glycoprotein; honey bee.
XX XX
XX XX Apis mellifera.
XX XX
XX XX Key Location/Qualifiers
FH Misc-difference 24
FT /label= Xaa= unknown
XX
XX JF2004008130-A.
XX XX
XX PD 15-JAN-2004.
XX XX
XX PF 10-JUN-2002; 2002JP-00168047.
XX XX
XX PR 10-JUN-2002; 2002JP-00168047.
XX XX
XX PA (POKK ) POLA CHEM IND INC.
XX XX
XX XX WPI; 2004-102884/11.
XX XX
XX XX Culture medium useful for culturing cells derived from animals, lacking
PT PT blood serum or blood serum components of animal and comprising

```

---

```

PT glycoprotein existing in royal jelly form, and which has a preset
PT molecular weight.
XX
XX Disclosure; SEQ ID NO 1; 8pp; Japanese.
XX
XX This invention relates to a novel culture medium lacking blood serum or a
XX blood serum component of an animal. Specifically, it refers to a serum-
XX free medium for cell culture that comprises a glycoprotein that exists in
XX royal jelly and forms a band in denaturing polyacrylamide gel
XX electrophoresis (PAGE) at a molecular weight of 57 kilo Daltons. The
XX present invention describes the serum as useful for the cultivation and
XX proliferation of cells derived from an animal such as a cancer cell,
XX hybridoma cell or stem cell. This is the honey bee peptide sequence of
XX the 57KD glycoprotein of the invention.
XX
XX Sequence 25 AA;
SQ

Query Match 73.2%; Score 101; DB 8; Length 25;
Best Local Similarity 87.0%; Pred. No. 2.4e-08;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NILRGESLNKSLPILHEWKFFDY 23
Db 1 NILRGESLLKLPILHEWKFFDY 23

RESULT 14
AAW94918
ID AAW94918 standard; protein; 768 AA.
XX
XX AC AAW94918;
XX XX
XX DT 04-MAY-1999 (first entry)
XX XX
XX DE Rat pheromone receptor Go-VN2.
XX XX
XX KW Pheromone receptor; signal transduction; fertility; behaviour;
XX XW reproduction; rodent; insect; rat.
XX XX
XX OS Rattus sp.
XX XX
XX PN WO9900422-A1.
XX XX
XX PD 07-JAN-1999.
XX XX
XX PF 30-JUN-1998; 98WO-US013680.
XX XX
XX PR 30-JUN-1997; 97US-0051284P.
XX XX
XX PA (HARD ) HARVARD COLLEGE.
XX XX
XX PI Buck L, Dulac C, Herrada G, Matsunami H;
XX XX
XX DR WPI; 1999-095684/08.
XX DR N-PSDB; AAX05818.
XX XX
XX XX New isolated pheromone receptor polypeptides - used to develop products
PT PT for controlling fertility and behaviour in vertebrates and invertebrates.
XX XX
XX PS Claim 33; Page 126-129; 308pp; English.
XX XX
XX CC The invention relates to polynucleotide sequences encoding mammalian
XX pheromone receptor polypeptides. The polypeptides are expressed in murine
XX and rat vomeronasal organ. The products can be used for modifying
XX pheromone activity, e.g. for decreasing pheromone receptor mediated
XX signal transduction. They can be used for controlling fertility and
XX behaviour in vertebrates and invertebrates. Compositions comprising the
XX polypeptides are particularly useful in e.g. controlling fertility in
XX livestock and controlling reproduction in rodents or insects by
XX interrupting the normal behaviours of rodents or insects that result in
XX reproduction. The present sequence represents a rat pheromone receptor Go
XX -VN2. The encoding cDNA is deposited under the Genbank accession number
XX AF016179

```



```

XX SQ Sequence 768 AA;
Query Match 40.2%; Score 55.5; DB 2; Length 768;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 11; Conservative 2; Mismatches 6; Indels 3; Gaps 1;

QY 3 LRGESLNKSLPILHEWKFFDYD 24
    || || : || ||| : ||
DB 259 LRGRDLNLLIP---EWKYFSYD 277

RESULT 15
ADL27606
ID ADL27606 standard; protein; 524 AA.
XX AC
XX AC ADL27606;
XX AC
XX AC 03-JUN-2004 (first entry)
XX AC Honey bee RJP70 mature protein, SEQ ID 3.
XX AC
XX AC Antiallergic; Antiasthmatic; Antiinflammatory; Dermatological; honey bee;
XX AC royal jelly protein; allergy; pollinosis; atopic dermatitis;
XX AC contact hypersensitivity; bronchial asthma; allergic rhinitis; RJP70.
XX AC
XX AC Apis mellifera.
XX AC
XX AC W02004019971-A1.
XX AC
XX AC 11-MAR-2004.
XX AC
XX AC 26-AUG-2003; 2003WO-JP010795.
XX AC
XX AC 29-AUG-2002; 2002JP-00252087.
XX AC
XX AC 30-JAN-2003; 2003JP-00022776.
XX AC
XX AC (HAYA/) HAYASHIBARA K.
XX AC
XX AC Okamoto I, Arai N, Kohno K, Kurimoto M, Sano O;
XX AC
XX AC WPI; 2004-248191/23.
XX AC
XX AC N-ESDB; ADL27608.
XX AC
XX AC Antiallergic agents for relieving symptoms accompanying e.g. pollinosis,
XX AC atopic dermatitis, contact hypersensitivity, bronchial asthma and
XX AC allergic rhinitis, comprise protein originated from royal jelly.
XX AC
XX AC Claim 1; SEQ ID NO 3; 78pp; Japanese.
XX AC
XX AC The present invention relates to novel antiallergic agents, which
XX AC comprise as an active ingredient, peptides ADL27604 and ADL27605, which
XX AC are N-terminal peptides of honey bee royal jelly proteins (RJP, ADL27606
XX AC and ADL27607). The agents can be used to relieve symptoms accompanying an
XX AC allergic disease e.g. pollinosis, atopic dermatitis, contact
XX AC hypersensitivity, bronchial asthma and allergic rhinitis, which are
XX AC applicable in foods, cosmetics and drugs.
XX AC
XX AC Sequence 524 AA;

Query Match 39.9%; Score 55; DB 8; Length 524;
Best Local Similarity 42.1%; Pred. No. 11;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 7 SLNKSPLILHEWKFFDYD 25
    :| | : : : ||| | : ||
DB 12 NLASHMKVIYEWKHIDFDF 30

Search completed: May 3, 2006, 19:01:25
Job time : 12.6224 secs

```

The page block (up to  
open) Xiong open

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2006, 19:01:48 ; Search time 1.87757 Seconds  
(without alignments)  
1281.133 Million cell updates/sec

Title: US-10-525-567-2

Sequence: 1 NLRGESLNKSLPILHEWKFFDYDF 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_80.\*  
1: PIR1.\*  
2: PIR2.\*  
3: PIR3.\*  
4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	73	52.9	464	2 S39194	royal jelly protei
2	55	39.9	467	2 S39193	royal jelly protei
3	53	38.4	404	2 H88650	protein C09G12.9
4	50.5	36.6	869	2 S49844	probable membrane
5	50.5	36.6	1115	2 T11614	probable poly(A)-s
6	50	36.2	546	2 A80088	probable flagellar
7	50	36.2	809	2 T41645	probable spindle p
8	49.5	35.9	161	2 G83914	hypothetical prote
9	49.5	35.9	1029	2 S57598	PP5 protein - yea
10	49	35.5	160	2 T15925	hypothetical prote
11	49	35.5	445	2 B37343	N-ethylamine ch
12	48	35.5	768	2 A11085	hypothetical prote
13	48	34.8	384	2 S51796	vasodilator-stimul
14	48	34.8	949	2 S54020	probable membrane
15	47	34.1	380	2 S51797	vasodilator-stimul
16	47	34.1	627	2 S53369	carnitine O-acetyl
17	46.5	33.7	357	2 F70193	hypothetical prote
18	46.5	33.7	439	2 T45798	hypothetical prote
19	46	33.3	271	2 H97230	alpha/beta superfa
20	46	33.3	327	2 F84610	hypothetical prote
21	46	33.3	393	2 JC5614	RNB6 protein - rat
22	46	33.3	526	2 A75581	flavin monooxime o
23	46	33.3	634	2 T27881	hypothetical prote
24	46	33.3	671	2 A30746	chemotaxis protein
25	46	33.3	671	2 A32959	chemotaxis protein
26	46	33.3	705	2 A86497	conserved hypothet
27	46	33.3	705	2 F72125	conserved hypothet
28	46	33.3	1021	2 T08601	hypothetical prote
29	46	33.3	1026	2 C97783	cell surface anti

30	45.5	33.0	553	2 I40807	cellulase (EC 3.2.
31	45.5	33.0	636	2 JC5874	cellulase (EC 3.2.
32	45	32.6	67	2 AC1954	hypothetical prote
33	45	32.6	166	2 A48465	ribosomal protein
34	45	32.6	256	2 H72349	conserved hypothet
35	45	32.6	425	2 E86737	malate transporter
36	45	32.6	425	2 B40576	cytochrome P450 AL
37	45	32.6	538	1 O4CKA4	cytochrome P450 52
38	45	32.6	582	2 T39931	probable transcript
39	45	32.6	627	2 D71574	hypothetical prote
40	45	32.6	628	2 B64534	hypothetical prote
41	45	32.6	772	2 T26313	hypothetical prote
42	44.5	32.2	473	2 T32038	hypothetical prote
43	44	31.9	139	2 A87442	Mut7/nudix family
44	44	31.9	219	2 JQ1606	multiple stimulus
45	44	31.9	240	2 H69846	hypothetical prote

ALIGNMENTS

RESULT 1

S39194

royal jelly protein RJP57-2 - honeybee

C;Species: Apis mellifera (honeybee)

C;Date: 05-Mar-1995 #sequence\_revision 12-May-1995 #text\_change 09-Jul-2004

C;Accession: S39194

R;Klaudiny, J.; Hanes, J.; Kulifajova, J.; Albert, S.; Simuth, J.

A;Submitted to the EMBL Data Library, September 1993

A;Description: Molecular cloning of two cDNAs from the head of the nurse honey bee (Apis

A;Reference number: S39193

A;Accession: S39194

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-464 <KLA>

A;Cross-references: UNIPROT:Q17061; UNIPARC:UPI000012F56A; EMBL:Z26319; NID:G4333530; PID

Query Match 52.9%; Score 73; DB 2; Length 464;

Best Local Similarity 47.6%; Pred. No. 0.004;

Matches 10; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Qy 5 GESLNKSLPILHEWKFFDYDF 25

Db 29 GKLTNTLNVTHKWKLYDYDF 49

RESULT 2

S39193

royal jelly protein RJP57-1 - honeybee

C;Species: Apis mellifera (honeybee)

C;Date: 05-Mar-1995 #sequence\_revision 12-May-1995 #text\_change 12-May-1995

C;Accession: S39193

R;Klaudiny, J.; Hanes, J.; Kulifajova, J.; Albert, S.; Simuth, J.

A;Submitted to the EMBL Data Library, September 1993

A;Description: Molecular cloning of two cDNAs from the head of the nurse honey bee (Apis

A;Reference number: S39193

A;Accession: S39193

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-467 <KLA>

A;Cross-references: UNIPARC:UPI0000179A48; EMBL:Z26318

Query Match 39.9%; Score 55; DB 2; Length 467;

Best Local Similarity 42.1%; Pred. No. 2.2;

Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 7 SLNKSPLILHEWKFFDYDF 25

Db 32 NLAHSMKVIYEWKIDDF 50

RESULT 3

H88650



A;Residues: 1-546 <KUR>  
A;Cross-references: UNIPROT:Q8Z108; UNIPARC:UPI000000DCD2A; GB:AL590842; PIDN:CA09565.1;  
C;Genetics:  
A;Gene: flif  
C;Superfamily: probable flagellar basal-body M ring protein

Query Match 36.2%; Score 50; DB 2; Length 546;  
Best Local Similarity 72.7%; Pred. No. 15;  
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 8 LNKSPLILHEW 18  
|||||:|:  
Db 360 LNKSPLVLEQW 370

RESULT 7  
T41645  
probable spindle pole body-associating protein - fission yeast (Schizosaccharomyces pombe)  
C;Species: Schizosaccharomyces pombe  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C;Accession: T41645  
R;Wood, V.; Rajandream, M.A.; Bartell, B.G.; Lucas, M.; Gaillardin, C.  
submitted to the EMBL Data Library, January 1999  
A;Reference number: 222006  
A;Accession: T41645  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-809 <WOO>  
A;Cross-references: UNIPROT:O94534; UNIPARC:UPI0000125794; EMBL:AL035247; PIDN:CAA22843.  
A;Experimental source: strain 972h-; cosmid c895  
C;Genetics:  
A;Gene: SPDB:SPCC895.07  
A;Map position: 3

Query Match 36.2%; Score 50; DB 2; Length 809;  
Best Local Similarity 50.0%; Pred. No. 24;  
Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 4 RGESLNKSLPILHEWKFFDYDF 25  
|:|:|:|:|:  
Db 303 RKEALDKLLPVLSPKIKDNDF 324

RESULT 8  
G83914  
hypothetical protein BH2119 [imported] - Bacillus halodurans (strain C-125)  
C;Species: Bacillus halodurans  
C;Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
C;Accession: G83914  
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira  
Nucleic Acids Res. 28, 4317-4331, 2000  
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
submitted to the EMBL Data Library, January 1999  
A;Reference number: A83650; MUID:20512582; PMID:11058132  
A;Accession: G83914  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-161 <STO>  
A;Cross-references: UNIPROT:Q9KB16; UNIPARC:UPI000000C3DBE1; GB:AP001514; GB:BA000004; NID  
A;Experimental source: strain C-125  
C;Genetics:  
A;Gene: BH2119

Query Match 35.9%; Score 49.5; DB 2; Length 161;  
Best Local Similarity 40.9%; Pred. No. 4.4;  
Matches 9; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

Qy 4 RGESLNKSLPILHEWKFFDYDF 25  
|:|:|:|:|:  
Db 99 KGEDIQKTVQDLHQW-LFDSY 119

RESULT 9  
S57598

PEP5 protein - yeast (Saccharomyces cerevisiae)  
N;Alternate names: ENL1 protein; protein YM9959.13; protein YMR231w  
C;Species: Saccharomyces cerevisiae  
C;Date: 19-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 09-Jul-2004  
C;Accession: S57598; S11188; S14460  
R;Skellton, J.; Churcher, C.M.  
submitted to the EMBL Data Library, June 1995  
A;Reference number: S57587  
A;Accession: S57598  
A;Molecule type: DNA  
A;Residues: 1-1029 <SKE>  
A;Cross-references: UNIPROT:PI2868; UNIPARC:UPI0000052FFA; EMBL:Z49939; NID:G887599; PID  
A;Experimental source: strain AB972  
R;Woolford, C.A.; Dixon, C.K.; Manolson, M.F.; Wright, R.; Jones, E.W.  
Genetics 125, 739-752, 1990  
A;Title: Isolation and characterization of PEP5, a gene essential for vacuolar biogenesis  
A;Reference number: S11188; MUID:90375022; PMID:2204580  
A;Accession: S11188  
A;Molecule type: DNA  
A;Residues: 1-138, 'K', 140-1023, 'E', 1024-1029 <WOO>  
A;Cross-references: UNIPARC:UPI0000179046; EMBL:X54466  
A;Experimental source: strain X2180-1B  
A;Note: the authors translated the codon GAA for residue 121 as Ile, GAC for residue 139  
R;Dulic, V.; Riezman, H.  
EMBO J. 8, 1349-1359, 1989  
A;Title: Characterization of the ENL1 gene required for vacuole biogenesis and gluconeog  
A;Reference number: S14460; MUID:89356602; PMID:2670552  
A;Accession: S14460  
A;Molecule type: DNA  
A;Residues: 1-36, 'R', 38-618, 'M', 620, 'FTVT', 624, 'H', 626-768, 'KQ', 771-1029 <DUL>  
A;Cross-references: UNIPARC:UPI0000168BFE; EMBL:X15355; NID:G3679; PIDN:CAA33413.1; PID:  
C;Genetics:  
A;Gene: SGD:PEP5; ENL1  
A;Cross-references: SGD:S0004844; MIPS:YMR231w  
A;Map position: 13R  
C;Function:  
A;Description: required for gluconeogenic growth; involved in both the internalization a  
C;Superfamily: Saccharomyces cerevisiae PEP5 protein  
C;Keywords: transmembrane protein; yeast vacuole  
F;139-155/Domain: transmembrane #status predicted <TM1>  
F;310-326/Domain: transmembrane #status predicted <TM2>  
F;454-470/Domain: transmembrane #status predicted <TM3>  
F;661-677/Domain: transmembrane #status predicted <TM4>

Query Match 35.9%; Score 49.5; DB 2; Length 1029;  
Best Local Similarity 40.0%; Pred. No. 38;  
Matches 10; Conservative 8; Mismatches 6; Indels 1; Gaps 1;

Qy 1 NILRGESLNKSLPILHEWKFFDYDF 25  
|:|:|:|:|:  
Db 52 NIIKVKLNQS-QVIEFQSPPHDF 75

RESULT 10  
T15925  
hypothetical protein EEED8.13 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
C;Accession: T15925  
R;Chissoe, S.  
submitted to the EMBL Data Library, July 1995  
A;Description: The sequence of C. elegans cosmid EEED8.  
A;Reference number: Z18428  
A;Accession: T15925  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-160 <CHI>  
A;Cross-references: UNIPROT:Q19068; UNIPARC:UPI0000007B4F; EMBL:U23484; NID:G733597; PII  
A;Experimental source: strain Bristol N2  
C;Genetics:  
A;Gene: CESP:EEED8.13  
A;Introns: 20/3

Query Match 35.5%; Score 49; DB 2; Length 160;  
Best Local Similarity 40.0%; Pred. No. 5.2;  
Matches 8: Conservative 5: Mismatches 7: Indels

QY 5 GESLNKSLPILHEWKFFDYD 24  
|:| | | :| | :| |  
Db 18 GORLNSQLNLHRSRKWIDFD 37

RESULT 11

B83743

N-ethylmaleimide chlorohydrolase BH0746 [imported] - Bacillus halodurans (strain C-125)

C;Species: Bacillus halodurans

C;Date: 01-Dec-2000 #text\_change 31-Dec-2004

C;Accession: B83743

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hiraoka, Y.

Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and reference number: A8350; MUID:20512582; PMID:11058132

A;Accession: B83743

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-445 <SFO>

A;Cross-references: UNIPROT:Q9KEV3; UNIPARC:UPI000000C3991; GB:AF001509; GB:BA000004; NID:1000000000

A;Experimental source: strain C-125

C;Genetics:

A;Gene: BH0746

Query Match 35.5%; Score 49; DB 2; Length 445;  
Best Local Similarity 47.1%; Pred. No. 17;  
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 6 ESLNKS LPILHEWKFFD 22  
: : : : :  
pb 165 OSI EKSIELLEWHSEF 181

RESULT 12  
AH1085  
hypothetical protein lmo0087 [imported] - *Listeria monocytogenes* (strain EGD-e)  
C:Species: *Listeria monocytogenes*  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: AH1085  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecher  
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkut, G.; Madueno, E.; Maitournam, A.; Ma  
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlund,  
A.; Title: Comparative genomics of *Listeria* species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AH1085  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-768 <GLA>  
A:Cross-references: UNIPROT:Q8YAN0; UNIPARC:UPI0000005487; GB:NC\_003210; PIDN:CAC98302.1  
A:Experimental source: strain EGD-e  
C:Genetics:  
A:Gene: lmo0087

Query Match 35.5%; Score 49; DB 2; Length 768;  
Best Local Similarity 34.8%; Pred. NO. 32;  
Matches 8: Conservative 6; Mismatches 9; Indels 0; Gaps 0;

Qy 1 NILRGESLNKSLPILHEWKFFDY 23  
| : : | | : : | |  
Db 540 NLKASQSVNKLQVQYDWSSADY 56

RESULT 13  
S51796  
vasodilator-stimulated phosphoprotein VASP - dog

C/Species: Canis lupus familiaris (dog)  
C/Date: 15-Jul-1995 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C/Accession: S51796  
R/Haffner, C.; Jarchau, T.; Reinhard, M.; Hoppe, J.; Lohmann, S.M.; Walter,  
EMBO J. 14, 19-27, 1995  
A/Title: Molecular cloning, structural analysis and functional expression o  
A/Reference number: S51796; MUID:95129547; PMID:7828592  
A/Accession: S51796  
A/Status: preliminary; nucleic acid sequence not shown  
A/Molecule type: mRNA  
A/Residues: 1-384 <HAF>  
A/Cross-references: UNIPROT:P50551; UNIPARC:UPI0000138121; EMBL:Z46388; NID  
C/Keywords: phosphoprotein

Query Match	34.8%	Score 48;	DB 2;	Length 384;
Best Local Similarity	38.9%	Pred. No. 21;		
Matches 7;	Conservative	5;	Mismatches	6;
			Indels	0;
			Gaps	0

Qy 2 ILRGESLNKSLPIHEWK 19  
| : | | | : : | : | :  
Db 66 IVRGIKYNOATPTFFHOWR 83

RESULT 14  
S54020  
probable membrane protein YMR019w - yeast (Saccharomycetes cerevisiae)  
N:Alternate names: hypothetical protein YM9711.08  
C:Species: Saccharomycetes cerevisiae  
C/Date: 08-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 05-Oct-2004  
C/Accession: S54020  
R:Lye, G.; Churcher, C.M.  
submitted to the EMBL Data Library, May 1995  
A:Reference number: S54014  
A:Accession: S54020  
A:Molecule type: DNA  
A:Residues: 1-949 <Lye>  
A:Cross-references: UNIPROT:P50104; UNIPARC:UPI00001360AE; EMBL:Z49211; NID  
A:Experimental source: strain A9972

Query Match	34.8%	Score	48;	DB	2;	Length	949;
Best Local Similarity	42.9%	Pred. No.	59;				
Matches	9;	Mismatches	5;				
Conservative	9;	Indels	7;				
Gaps	0;						

Qy 4 RGESLNKSLPILHEWKFFDYD 24  
          :-: : -| | | | :-|  
Db 369 KGDFRSEKNIRHESKEFFED 389

RESULT 15  
S51797  
vasodilator-stimulated phosphoprotein - human  
C:Species: Homo sapiens (man)  
C:Date: 15-Jul-1995 #sequence\_revision 09-Mar-1996 #text\_change 09-Jul-2004  
C:Accession: S51797; S55526  
F:Haefliger, C.; Jarchau, T.; Reinhard, M.; Hoppe, J.; Lohmann, S.M.; Walter,  
EMBO J. 14, 19-27, 1995  
A:Title: Molecular cloning, structural analysis and functional expression c  
A:Reference number: S51796; MUID:95129547; PMID:7828592

RESULT 13

S51796

vasodilator-stimulated phosphoprotein VASP - dog

---

A:Residues: 1-380 <HAF1>  
A:Cross-references: UNIPROT:P50552; UNIPARC:UPI00000001C69; EMBL:Z46389; NID:9624963; PID:  
A:Accession: S55526  
A:Molecule type: protein  
A:Residues: 11-32;87-96;140-154;255-282;297-303;305-322 <HAF2>

A;Cross-references: UNIPARC:UPI000017C415; UNIPARC:UPI000017C416; UNIPARC:UPI000017C417;  
C;keywords: phosphoprotein

Query Match 34.1%; Score 47; DB 2; Length 380;  
Best Local Similarity 38.9%; Pred. No. 29;  
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 ILRGESLNKSLPILHEWK 19  
|:| |::| |:|  
Db 66 IVRGVKNQATPNFQWR 83

Search completed: May 3, 2006, 19:10:18  
Job time : 2.87757 secs

This Page Blank (uspo)



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2006, 18:54:38 ; Search time 11.5741 Seconds  
(without alignments)  
1523.941 Million cell updates/sec

Title: US-10-525-567-2  
Perfect score: 138  
Sequence: 1 NLRGSLNKSUPILHEWKEFFDYDF 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot 05.80.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	138	100.0	432	1 MRJPI_APIME	O18330 apis mellif
2	138	100.0	432	2 Q548D6_APIME	Q548D6 apis mellif
3	124	89.9	222	2 Q51224_APICE	Q51224 apis cerana
4	124	89.9	433	2 Q5VLE2_APICE	Q5VLE2 apis cerana
5	124	89.9	433	2 Q8MMJ1_APICE	Q8MMJ1 apis cerana
6	84	60.9	452	1 MRJPI_APIME	O77061 apis mellif
7	76.5	55.4	443	2 Q5IMJ9_APIME	Q5IMJ9 apis mellif
8	75.5	54.7	220	2 Q51223_APICE	Q51223 apis cerana
9	74	53.6	463	2 Q8MMJ0_APICE	Q8MMJ0 apis cerana
10	74	53.6	463	2 Q5OH66_APICE	Q5OH66 apis cerana
11	74	53.6	468	2 Q5VK55_APICE	Q5VK55 apis cerana
12	73	52.9	464	1 MRJPI_APIME	O17061 apis mellif
13	71	51.4	437	2 Q8W3E3_APIME	Q8W3E3 apis mellif
14	70	50.7	416	2 Q5TGR0_APIME	Q5TGR0 apis mellif
15	69	50.0	222	2 Q51222_APICE	Q51222 apis cerana
16	69	50.0	485	2 Q6QNT6_APICE	Q6QNT6 apis cerana
17	69	50.0	579	2 Q6QNT5_APICE	Q6QNT5 apis cerana
18	69	50.0	598	2 Q5VK56_APICE	Q5VK56 apis cerana
19	68	49.3	598	1 MRJPI_APIME	O97432 apis mellif
20	60	43.5	274	2 Q8TKD5_METAC	Q8TKD5 methanosarc
21	57	41.3	131	2 Q9D261_MOUSE	Q9D261 mus musculus
22	55.5	40.2	768	2 Q35266_RAT	Q35266 rattus norv
23	55	39.9	375	1 VASP_MOUSE	P70460 mus musculus
24	55	39.9	375	2 Q91VD2_MOUSE	Q91VD2 mus musculus
25	55	39.9	544	1 MRJPI_APIME	O17060 apis mellif
26	54.5	39.5	509	1 MATK_SCHTR	Q95E84 schlumberge
27	53	38.4	310	2 Q8IA99_CABEL	Q8IA99 caenorhabdi
28	53	38.4	425	2 Q76258_CABEL	Q76258 caenorhabdi
29	53	38.4	454	2 Q8IAA0_CABEL	Q8IAA0 caenorhabdi
30	53	38.4	468	2 Q8IT88_CABEL	Q8IT88 caenorhabdi
31	52	37.7	536	2 Q6C7B7_YARLLI	Q6C7B7 yarrowia li

32	52	37.7	712	2	Q6BU38_DEBHA	Q6BU38 debaryomyce
33	52	37.7	1072	2	Q4WXA3_ASPPU	Q4WXA3 aspergillus
34	51	37.0	281	2	Q6G7T5_STAAS	Q6G7T5 staphylococ
35	51	37.0	281	2	Q5HEH8_STAAC	Q5HEH8 staphylococ
36	51	37.0	281	2	Q8NVL6_STAAM	Q8NVL6 staphylococ
37	51	37.0	495	2	Q8D7I4_VIBVU	Q8D7I4 vibrio vuln
38	51	37.0	1250	2	Q5B8W8_EMENI	Q5B8W8 aspergillus
39	50.5	36.6	869	1	YD95_YEAST	P38966 saccharomyc
40	50.5	36.6	1115	1	YAA4_SCHPO	Q9798 schizosacch
41	50	36.2	240	2	Q82S24_NITEU	Q82S24 nitrosomona
42	50	36.2	423	2	Q4ZJX1_APIME	Q4ZJX1 apis mellif
43	50	36.2	546	2	Q8Z108_YERPE	Q8Z108 yersinia pe
44	50	36.2	546	2	Q66B80_YERPS	Q66B80 yersinia ps
45	50	36.2	809	1	ALP14_SCHPO	Q94534 schizosacch

## ALIGNMENTS

### RESULT 1

ID	MRJPI_APIME	STANDARD;	PRT;	432 AA.
AC	O18330;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	10-MAY-2005 (Rel. 47, Last annotation update)			
DE	Major royal jelly protein 1 precursor (MRJP-1) (Bee-milk protein).			
GN	Names=MRJPI;			
OS	Apis mellifera (Honeybee).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Hymenoptera; Apoidea; Apoidea;			
OC	Apidae; Apis.			
OX	NCBI_TaxID=7460;			
RL	[1]			
RP	NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 39-48; 63-71 AND 361-380.			
RC	TISSUE=Hypopharyngeal gland;			
RX	MEDLINE=98055707; PubMed=9395329;			
RA	Ohashi K., Natori S., Kubo T.;			
RT	"Change in the mode of gene expression of the hypopharyngeal gland cells with an age-dependent role change of the worker honeybee Apis mellifera L.";			
RT	Eur. J. Biochem. 249:797-802(1997).			
RL	[2]			
RP	NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 20-30.			
RC	TISSUE=Head;			
RX	MEDLINE=99007754; PubMed=9791542; DOI=10.1007/s000180050229;			
RA	Schmitzova J., Klaudivny J., Albert S., Schroeder W., Schreckengost W., Hanes J., Judova J., Simuth J.;			
RT	"A family of major royal jelly proteins of the honeybee Apis mellifera L.";			
RL	Cell. Mol. Life Sci. 54:1020-1030(1998).			
CC	-1- FUNCTION: May play an important role in honeybee nutrition. Most abundant protein found in the royal jelly which is the food of the queen honey bee larva. The royal jelly determines the development of the young larvae and is responsible for the high reproductive ability of the honeybee queen.			
CC	-1- SUBCELLULAR LOCATION: Secreted.			
CC	-1- TISSUE SPECIFICITY: Found in the hypopharyngeal glands of the worker honeybee.			
CC	-1- DEVELOPMENTAL STAGE: Produced in the cephalic glands of both the nurse bee and the forager bee. This bee milk protein changes to alpha-glucosidase in accordance with the age-dependent role change of the worker bee.			
CC	-1- SIMILARITY: Belongs to the major royal jelly protein family.			

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

EMBL; D79207; BAA23639.1; -; mRNA.  
EMBL; AF000633; AAC61895.1; -; mRNA.



```

DR Pfam; PF03022; MRJP; 1.
DR PRINTS; PR01366; ROYALJELLY.
KW Direct protein sequencing; Glycoprotein; Signal.
FT SIGNAL 1 17
FT CHAIN 18 452 Major royal jelly protein 2.
FT CARBOHYD 145 145 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 178 178 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 452 AA; 51074 MW; 25E5B621288FE189 CRC64;

Query Match 60.9%; Score 84; DB 1; Length 452;
Best Local Similarity 58.4%; Pred. No. 0.00057;
Matches 13; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 7 SLNKSPLILHEWKFFDYDF 25
   :||| :|||:||||
Db 27 NLEKSLNVIHEWKFFDYDF 45

RESULT 7
Q6IMJ9 APIME PRELIMINARY; PRT; 443 AA.
ID Q6IMJ9;
AC Q6IMJ9;
DT 05-JUL-2004 (T-REMBLrel. 27, Created)
DT 05-JUL-2004 (T-REMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (T-REMBLrel. 27, Last annotation update)
DE Major royal jelly protein 7.
DE Name=MRJP7;
OS Apis mellifera (Honeybee).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
OC Apidae; Apis.
OC NCBI_TaxID=7460;
[1]
RN NUCLEOTIDE SEQUENCE.
RX PubMed=15037093; DOI=10.1016/j.jinsphys.2003.09.008;
RA Albert S. Klaudiny J.;
RT "The MRJP/YELLOW protein family of Apis mellifera: identification of
RL J. Insect Physiol. 50:51-59(2004).
RC -I- MISCELLANEOUS: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ third party annotation (TPA) entry.
DR EMBL; BK001420; DAA01512.1; -; mRNA.
DR Ensembl; ENSAPG0000007331; Apis mellifera.
DR InterPro; IPR003534; Royaljelly.
DR Pfam; PF03022; MRJP; 1.
DR PRINTS; PR01366; ROYALJELLY.
DR SEQUENCE 443 AA; 50541 MW; 8916272B4A4DB421 CRC64;

Query Match 55.4%; Score 76.5; DB 2; Length 443;
Best Local Similarity 55.6%; Pred. No. 0.0078;
Matches 15; Conservative 3; Mismatches 6; Indels 3; Gaps 1

QY 2 ILRGES---LNKSLPILHEWKFFDYDF 25
   ||| ||| ||| :|||:||||
Db 19 ILRENSARNLKNLSKVMHWEKYIDYDF 45

RESULT 8
Q51223 APICE
ID Q51223 APICE PRELIMINARY; PRT; 220 AA.
AC Q51223;
DT 10-MAY-2005 (T-REMBLrel. 30, Created)
DT 10-MAY-2005 (T-REMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (T-REMBLrel. 30, Last annotation update)
DE Major royal jelly protein 7 (Fragment).
DE Name=MRJP7;
OS Apis cerana (Indian honeybee).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
OC Apidae; Apis.
OC NCBI_TaxID=7461;
[1]
RN NUCLEOTIDE SEQUENCE.

```

```
RA Su S., Albert S., Chen S., Zhong B.;
RT "Molecular cloning and analysis of four cDNAs from the heads of Apis
RT cerana cerana nurse honeybees coding for major royal jelly proteins.";
RL Apidologie 36:389-401(2005).
DR EMBL: AY862496; AAW51951.1; -: Genomic_DNA.
DR InterPro: IPR003534; Royaljelly.
DR Pfam: PF03022; MRJP; 1.
DR PRINTS; PRO1366; ROYALJELLY.
FT NON TER 220 220
FT NON TER 1
SQ SEQUENCE 220 AA; 24927 MW; 7EC00E3ETAAA6822 CRC64;

Query Match 54.7%; Score 75.5; DB 2; Length 220;
Best Local Similarity 55.6%; Pred. No. 0.0051;
Matches 15; Conservative 4; Mismatches 5; Indels 3; Gaps 1;

QY 2 ILRGES---LNKSLPILHEWKFFDYDF 25
Db 10 IVRKSSARNLENSLVNVIHEWKYIDYDF 36

RESULT 9
Q8MMJ0.APICE PRELIMINARY; PRT; 463 AA.
AC Q8MMJ0;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Major royal jelly protein MRJP2 precursor.
GN Name=MRJP2;
OS Apis cerana (Indian honeybee).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
OC Apidae; Apis.
OX NCBI_TaxID=7461;
RN [1];
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Nurse heads;
RA Sittipraneed S., Injongjirak C.;
RT "Molecular Cloning of Major Royal Jelly Protein (MRJP2) cDNA from Apis
RT cerana in Thailand.";
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF525777; AAM89282.2; -: mRNA.
DR InterPro: IPR003534; Royaljelly.
DR Pfam: PF03022; MRJP; 1.
DR PRINTS; PRO1366; ROYALJELLY.
KW Signal.
FT SIGNAL 1 17 Potential.
FT SIGNAL 463 AA; 52490 MW; 29D92E3749B15992 CRC64;

Query Match 53.6%; Score 74; DB 2; Length 463;
Best Local Similarity 55.0%; Pred. No. 0.02;
Matches 11; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 6 ESLNKSILPILHEWKFFDYDF 25
Db 26 KNLNSLVNVIHEWKYIDYDF 45

RESULT 10
Q50H66.APICE PRELIMINARY; PRT; 463 AA.
AC Q50H66;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Major royal jelly protein 2.
GN Name=mrj1;
OS Apis cerana (Indian honeybee).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
OC Apidae; Apis.
OX NCBI_TaxID=7461;
```

```
RN NUCLEOTIDE SEQUENCE.
RP Injongjirak C., Klinbunga S., Sittipraneed S.;
RT "Cloning, Expression and Genomic Organization of Genes Encoding Major
RT Royal Jelly Protein 1 and 2 of the Honeybee (Apis cerana).";
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY515689; AAS88557.1; -: Genomic_DNA.
SQ SEQUENCE 463 AA; 52462 MW; 930C91C8CC595935 CRC64;

Query Match 53.6%; Score 74; DB 2; Length 463;
Best Local Similarity 55.0%; Pred. No. 0.02;
Matches 11; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 6 ESLNKSILPILHEWKFFDYDF 25
Db 26 KNLNSLVNVIHEWKYIDYDF 45

RESULT 11
Q5VK55.APICC PRELIMINARY; PRT; 468 AA.
AC Q5VK55;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Major royal jelly protein MRJP2.
OS Apis cerana cerana (Oriental honeybee).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
OC Apidae; Apis.
OX NCBI_TaxID=94128;
RN [1];
RP NUCLEOTIDE SEQUENCE.
RA Su S., Chen S., Albert S.;
RT "Molecular cloning of MRJP2 cDNA from Apis cerana cerana in China.";
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY392758; AAR83083.1; -: mRNA.
DR InterPro: IPR003534; Royaljelly.
DR Pfam: PF03022; MRJP; 1.
DR PRINTS; PRO1366; ROYALJELLY.
SQ SEQUENCE 468 AA; 53060 MW; 5B2A6AEP4C530C84 CRC64;

Query Match 53.6%; Score 74; DB 2; Length 468;
Best Local Similarity 55.0%; Pred. No. 0.02;
Matches 11; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 6 ESLNKSILPILHEWKFFDYDF 25
Db 26 KNLNSLVNVIHEWKYIDYDF 45

RESULT 12
MRJP4.APIME STANDARD; PRT; 464 AA.
AC Q17061;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Major royal jelly protein 4 precursor (MRJP-4) (Bee-milk protein)
DE (Royal jelly protein RJP57-2).
GN Name=MRJP4;
OS Apis mellifera (Honeybee).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
OC Apidae; Apis.
OX NCBI_TaxID=7460;
RN [1];
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Head;
RA Klaudiny J., Hanes J., Kulifajova J., Albert S., Simuth J.;
RT "Molecular cloning of two cDNAs from the head of the nurse honey bee
RT (Apis mellifera L.) for coding related proteins of royal jelly.";
RL J. Apicultural Res. 33:105-111 (1994).
```



```
RL Apidologie 36:389-401(2005).
DR ENBL; AY862497; AAW51952.1; -; Genomic_DNA.
DR InterPro; IPR003534; RoyalJelly.
DR Pfam; PF03022; MRJP; 1.
DR PRINTS; PR01366; ROYALJELLY.
FT NON_TER 222 222
SQ SEQUENCE 222 AA; 24877 MW; B37705F4570C549F CRC64;

Query Match      50.0%; Score 69; DB 2; Length 222;
Best Local Similarity 52.6%; Pred.No; 0.05;
Matches 10; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy      7 SLNKSFLPILHEWKFFDYDF 25
       :| | : :| | | :| | |
Db      21 NLANSNMVNIHEWKVLDYDF 39

Search completed: May 3, 2006, 19:09:00
Job time : 13.5741 secs
```

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2006, 19:09:24 ; Search time 2.70062 Seconds  
(without alignments)  
765.341 Million cell updates/sec

Title: US-10-525-567-2

Perfect score: 138

Sequence: 1 NILRGESLNKSLPILHEWKFFDYDF 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:  
1: /cgm2\_6/ptodata/1/iaa/5 COMB.pap.\*  
2: /cgm2\_6/ptodata/1/iaa/6 COMB.pap.\*  
3: /cgm2\_6/ptodata/1/iaa/H COMB.pap.\*  
4: /cgm2\_6/ptodata/1/iaa/PCTUS COMB.pap.\*  
5: /cgm2\_6/ptodata/1/iaa/RE COMB.pap.\*  
6: /cgm2\_6/ptodata/1/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	36.2	696	2	US-09-248-796A-16205
2	49.5	35.9	274	2	US-08-256-847C-4
3	49.5	35.9	411	2	US-09-328-352-5039
4	49.5	35.9	610	2	US-09-328-352-7804
5	48	34.8	469	2	US-09-248-796A-16207
6	48	34.8	508	2	US-09-248-796A-16204
7	47.5	34.4	204	2	US-09-914-098-40
8	47	34.1	116	2	US-09-377-285B-69
9	47	34.1	116	2	US-10-192-381-69
10	47	34.1	378	1	US-08-986-217-6
11	47	34.1	380	1	US-09-026-587-4
12	47	34.1	380	1	US-09-227-420-4
13	47	34.1	380	2	US-09-387-811-4
14	47	34.1	380	2	US-09-823-240A-10
15	47	34.1	381	2	US-09-949-016-7291
16	47	34.1	635	2	US-09-417-197-125
17	46.5	33.7	444	2	US-10-104-047-2354
18	46	33.3	115	2	US-09-377-285B-67
19	46	33.3	115	2	US-10-192-381-67
20	46	33.3	232	2	US-09-328-352-5918
21	46	33.3	314	2	US-09-540-236-2699
22	46	33.3	393	1	US-09-026-587-3
23	46	33.3	393	1	US-09-227-420-3
24	46	33.3	393	2	US-09-387-811-3
25	46	33.3	393	2	US-09-823-240A-11
26	46	33.3	418	1	US-09-026-587-1
27	46	33.3	418	1	US-09-227-420-1

28	46	33.3	418	2	US-09-387-811-1	Sequence 1, Appl
29	46	33.3	705	2	US-09-198-452A-68	Sequence 68, Appl
30	46	33.3	705	2	US-09-438-185A-50	Sequence 50, Appl
31	45	32.6	166	2	US-09-732-210-702	Sequence 702, Appl
32	44.5	32.2	1786	1	US-08-477-451-16	Sequence 16, Appl
33	44	31.9	240	2	US-09-355-166-19	Sequence 19, Appl
34	44	31.9	305	2	US-10-012-231A-273	Sequence 273, Appl
35	44	31.9	305	2	US-10-015-389A-273	Sequence 273, Appl
36	44	31.9	305	2	US-10-006-768A-273	Sequence 273, Appl
37	44	31.9	305	2	US-10-015-671A-273	Sequence 273, Appl
38	44	31.9	305	2	US-10-015-393A-273	Sequence 273, Appl
39	44	31.9	305	2	US-10-011-833A-273	Sequence 273, Appl
40	44	31.9	305	2	US-10-006-041A-273	Sequence 273, Appl
41	44	31.9	305	2	US-10-012-064A-273	Sequence 273, Appl
42	44	31.9	333	2	US-09-107-532A-4698	Sequence 4698, Appl
43	44	31.9	389	2	US-09-270-767-46824	Sequence 46824, Appl
44	44	31.9	481	2	US-09-215-694-17	Sequence 17, Appl
45	44	31.9	481	2	US-10-109-310-17	Sequence 17, Appl

## ALIGNMENTS

RESULT 1  
US-09-248-796A-16205  
; Sequence 16205, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 16205  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-09-248-796A-16205

Query Match 36.2%; Score 50; DB 2; Length 696;  
Best Local Similarity 38.7%; Pred. No. 11;  
Matches 12; Conservative 6; Mismatches 7; Indels 6; Gaps 1;

Qy 1 NILRGESLNKSL-----PILHEWKFFDYDF 25  
Db 189 SILPGLANKYLTMLARKEIINGWKISSYNF 219

RESULT 2  
US-08-256-847C-4  
; Sequence 4, Application US/08256847C  
; Patent No. 6403099  
; GENERAL INFORMATION:  
; APPLICANT: Rappuoli, Rino  
; APPLICANT: Constantino, Paolo  
; APPLICANT: No. 6403099elli, Francesco  
; TITLE OF INVENTION: Conjugates Formed From Heat Shock Proteins And Oligo or Polysaccch

; FILE REFERENCE: CHIR-0042  
; CURRENT APPLICATION NUMBER: US/08/256,847C  
; CURRENT FILING DATE: 1994-11-01  
; PRIOR APPLICATION NUMBER: PCT/EP93/00516  
; PRIOR FILING DATE: 1993-03-08  
; PRIOR APPLICATION NUMBER: F192A000058  
; PRIOR FILING DATE: 1992-03-06  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4

```
; LENGTH: 274
; TYPE: PRT
; ORGANISM: M. leprae
US-08-256-847C-4

Query Match      35.9%; Score 49.5; DB 2; Length 274;
Best Local Similarity 45.5%; Pred. No. 4.5;
Matches 10; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

Qy 4 RGSLSKSLPILHEW-KFFDYD 24
    ||| ||| : : : ||| :
Db 54 RGERLSALKVLEKWTNIDYE 75

RESULT 3
US-09-328-352-5039
; Sequence 5039, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5039
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5039

Query Match      35.9%; Score 49.5; DB 2; Length 411;
Best Local Similarity 52.2%; Pred. No. 7.2;
Matches 12; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

Qy 1 NILRGESLNKSLPILHEWKFFDY 23
    | : | | | | | | | | | | | |
Db 252 NVLRRSLNTHLEALHQQ-FDY 273

RESULT 4
US-09-328-352-7804
; Sequence 7804, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7804
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7804

Query Match      35.9%; Score 49.5; DB 2; Length 610;
Best Local Similarity 62.5%; Pred. No. 12;
Matches 10; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

Qy 10 KSLPILH-EWKFFDYD 24
    | | | | | | | | | |
Db 588 KGPLHAEWYFDDID 603

RESULT 5
US-09-248-796A-16207
; Sequence 16207, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
```

```
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16207
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-16207

Query Match      34.8%; Score 48; DB 2; Length 469;
Best Local Similarity 35.5%; Pred. No. 15;
Matches 11; Conservative 7; Mismatches 7; Indels 6; Gaps 1;

Qy 1 NILRGESLNKSL-----PILHEWKFFDYDF 25
    : || | : ||| : | : : | : |
Db 331 SILPGLALNKLTLPAPKEVINGWKISGYNF 361

RESULT 6
US-09-248-796A-16204
; Sequence 16204, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16204
; LENGTH: 508
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (487),(503),(504),(506)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown
US-09-248-796A-16204

Query Match      34.8%; Score 48; DB 2; Length 508;
Best Local Similarity 35.5%; Pred. No. 17;
Matches 11; Conservative 7; Mismatches 7; Indels 6; Gaps 1;

Qy 1 NILRGESLNKSL-----PILHEWKFFDYDF 25
    : || | : ||| : | : : | : |
Db 301 SILPGLALNKLTLPAPKEVINGWKISGYNF 331

RESULT 7
US-09-914-098-40
; Sequence 40, Application US/09914098
; Patent No. 6855863
; GENERAL INFORMATION:
; APPLICANT: E. I. du Pont de Nemours and Company
; TITLE OF INVENTION: Lysophosphatidic Acid Acetyltransferases
; FILE REFERENCE: BB1332
; CURRENT APPLICATION NUMBER: US/09/914,098
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/121,119
; PRIOR FILING DATE: 1999-02-22
```



```
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 40
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (68)
; OTHER INFORMATION: Xaa = any amino acid
US-09-914-098-40

Query Match          34.1%; Score 47.5; DB 2; Length 204;
Best Local Similarity 47.1%; Pred. No. 6.8;
Matches 8; Conservative 5; Mismatches 1; Indels 1; Gaps 1;

Db 168 ALN-ALPTYYSWRFPDY 183
      ||| :|| : |||
      7 SLNKSPLILHEWKFFDY 23

RESULT 8
US-09-377-285B-69
; Sequence 69, Application US/09377285B
; Patent No. 6720175
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: WORLEY, Paul
; APPLICANT: TU, Jian
; APPLICANT: XIAO, Bo
; APPLICANT: LEAHY, Daniel
; APPLICANT: BENEKEN, Jutta
; APPLICANT: LANAHAN, Anthony
; TITLE OF INVENTION: NUCLEIC ACID MOLECULE ENCODING HOMER 1b PROTEIN (AS AMENDED)
; FILE REFERENCE: JHU1580-4
; CURRENT APPLICATION NUMBER: US/09/377,285B
; CURRENT FILING DATE: 1999-08-18
; PRIOR APPLICATION NUMBER: US 60/138,426
; PRIOR FILING DATE: 1999-06-10
; PRIOR APPLICATION NUMBER: US 60/138,493
; PRIOR FILING DATE: 1999-06-10
; PRIOR APPLICATION NUMBER: US 60/138,494
; PRIOR FILING DATE: 1999-06-10
; PRIOR APPLICATION NUMBER: US 60/097,334
; PRIOR FILING DATE: 1998-08-18
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 69
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-377-285B-69

Query Match          34.1%; Score 47; DB 2; Length 116;
Best Local Similarity 38.9%; Pred. No. 4.2;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ILRGESLNKSLPILHEWK 19
      ||| :|| : |||
      66 IVRGVKYNQATPNFQWR 83

Db 66 IVRGVKYNQATPNFQWR 83

RESULT 10
US-08-986-217-6
; Sequence 6, Application US/08986217
; Patent No. 5914249
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: CELL-CYCLE PHOSPHOPROTEINS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/986,217
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0452 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 378 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

```
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 40
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (68)
; OTHER INFORMATION: Xaa = any amino acid
US-09-914-098-40

Query Match          34.4%; Score 47.5; DB 2; Length 204;
Best Local Similarity 47.1%; Pred. No. 6.8;
Matches 8; Conservative 5; Mismatches 1; Indels 1; Gaps 1;

Db 168 ALN-ALPTYYSWRFPDY 183
      ||| :|| : |||
      7 SLNKSPLILHEWKFFDY 23

RESULT 8
US-09-377-285B-69
; Sequence 69, Application US/09377285B
; Patent No. 6720175
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: WORLEY, Paul
; APPLICANT: TU, Jian
; APPLICANT: XIAO, Bo
; APPLICANT: LEAHY, Daniel
; APPLICANT: BENEKEN, Jutta
; APPLICANT: LANAHAN, Anthony
; TITLE OF INVENTION: NUCLEIC ACID MOLECULE ENCODING HOMER 1b PROTEIN (AS AMENDED)
; FILE REFERENCE: JHU1580-4
; CURRENT APPLICATION NUMBER: US/09/377,285B
; CURRENT FILING DATE: 1999-08-18
; PRIOR APPLICATION NUMBER: US 60/138,426
; PRIOR FILING DATE: 1999-06-10
; PRIOR APPLICATION NUMBER: US 60/138,493
; PRIOR FILING DATE: 1999-06-10
; PRIOR APPLICATION NUMBER: US 60/138,494
; PRIOR FILING DATE: 1999-06-10
; PRIOR APPLICATION NUMBER: US 60/097,334
; PRIOR FILING DATE: 1998-08-18
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 69
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-377-285B-69

Query Match          34.1%; Score 47; DB 2; Length 116;
Best Local Similarity 38.9%; Pred. No. 4.2;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ILRGESLNKSLPILHEWK 19
      ||| :|| : |||
      66 IVRGVKYNQATPNFQWR 83

Db 66 IVRGVKYNQATPNFQWR 83

RESULT 9
US-10-192-381-69
; Sequence 69, Application US/10192381
; Patent No. 6864083
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: WORLEY, Paul
; APPLICANT: TU, Jian
; APPLICANT: XIAO, Bo
; APPLICANT: LEAHY, Daniel
; APPLICANT: BENEKEN, Jutta
```

```
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1617319
US-08-986-217-6

Query Match          34.1%; Score 47; DB 1; Length 378;
Best Local Similarity 38.9%; Pred. No. 17;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ILRGESLNKSLPILHEWK 19
Db 64 IVRGVKYNQATPNFQWR 81

RESULT 11
US-09-026-587-4
; Sequence 4, Application US/09026587
; Patent No. 5912128
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN ENA/VASP-LIKE PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/026,587
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0471 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 624964
US-09-026-587-4

Query Match          34.1%; Score 47; DB 1; Length 380;
Best Local Similarity 38.9%; Pred. No. 17;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ILRGESLNKSLPILHEWK 19
Db 66 IVRGVKYNQATPNFQWR 83

RESULT 12
US-09-026-587-4
; Sequence 4, Application US/09026587
; Patent No. 5912128
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN ENA/VASP-LIKE PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/026,587
; FILING DATE: Filled Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0471 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 624964
US-09-026-587-4

Query Match          34.1%; Score 47; DB 1; Length 380;
Best Local Similarity 38.9%; Pred. No. 17;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ILRGESLNKSLPILHEWK 19
Db 66 IVRGVKYNQATPNFQWR 83

RESULT 13
US-09-387-811-4
; Sequence 4, Application US/09387811
; Patent No. 6645499
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN ENA/VASP-LIKE PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/227,420
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/026,587
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0471 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 624964
US-09-227-420-4

Query Match          34.1%; Score 47; DB 1; Length 380;
Best Local Similarity 38.9%; Pred. No. 17;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ILRGESLNKSLPILHEWK 19
Db 66 IVRGVKYNQATPNFQWR 83
```

```

;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/387,811
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/026,587
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0471 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 624964
;
US-09-387-811-4
;
Query Match 34.1%; Score 47; DB 2; Length 380;
Best Local Similarity 38.9%; Pred. No. 17;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ILRGESLNKSLPILHEWK 19
|:|:| |:| |:| |:| |:|
Db 66 IVRGVKNQATPNFHQWR 83

RESULT 14
US-09-823-240A-10
; Sequence 10, Application US/09823240A
; Patent No. 6716597
; GENERAL INFORMATION:
; APPLICANT: Frank B. Gertler
; APPLICANT: James B. Bear
; APPLICANT: Jurgen Wehland
; APPLICANT: Joseph Loureiro
; TITLE OF INVENTION: Methods and Products for Regulating Cell
; FILE REFERENCE: M00656.70064.US
; CURRENT APPLICATION NUMBER: US/09/823,240A
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 60/194,564
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-823-240A-10

Query Match 34.1%; Score 47; DB 2; Length 380;
Best Local Similarity 38.9%; Pred. No. 17;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ILRGESLNKSLPILHEWK 19
|:|:| |:| |:| |:| |:|
Db 66 IVRGVKNQATPNFHQWR 83

RESULT 15
US-09-949-016-7291
; Sequence 7291, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7291
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7291

Query Match 34.1%; Score 47; DB 2; Length 381;
Best Local Similarity 38.9%; Pred. No. 17;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ILRGESLNKSLPILHEWK 19
|:|:| |:| |:| |:| |:|
Db 67 IVRGVKNQATPNFHQWR 84

Search completed: May 3, 2006, 19:12:05
Job time : 3.70062 secs
```

Page Mark (tuple)

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2006, 19:10:39 ; Search time 8.77058 Seconds  
(without alignments)  
1190.997 Million cell updates/sec

Title: US-10-525-567-2

Perfect score: 138  
Sequence: 1 NIURGESLNKSLPILHWEKFFDYDF 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA Main:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	39.9	79	4	US-10-424-599-242162
2	52	37.7	42	3	US-09-764-877-1036
3	52	37.7	42	4	US-10-242-515-1036
4	52	37.7	168	4	US-10-424-599-180168
5	51	37.0	533	4	US-10-425-115-266449
6	49.5	35.9	255	4	US-10-282-122A-44667
7	49	35.5	142	4	US-10-418-861B-46
8	49	35.5	768	4	US-10-282-122A-60551
9	48	34.8	152	4	US-10-437-963-106441
10	48	34.8	439	4	US-10-424-599-190654
11	48	34.8	445	4	US-10-425-114-37354
12	48	34.8	541	4	US-10-437-963-106443
13	48	34.8	904	4	US-10-032-585-7917
14	48	34.8	949	4	US-10-149-310-206
15	47	34.1	46	4	US-10-425-115-201291
16	47	34.1	85	4	US-10-193-144-69
17	47	34.1	85	4	US-10-345-072-69
18	47	34.1	116	4	US-10-192-381-69
19	47	34.1	116	6	US-11-008-889-69
20	47	34.1	242	4	US-10-425-114-45051
21	47	34.1	242	4	US-10-425-114-53535
22	47	34.1	380	3	US-09-823-240-10
23	47	34.1	380	4	US-10-311-093-1
24	47	34.1	380	4	US-10-618-538-4
25	47	34.1	571	4	US-10-424-599-144073
26	47	34.1	635	4	US-10-072-036-125
27	47	34.1	1088	4	US-10-437-963-204592

28	47	34.1	1628	4	US-10-437-963-170587	Sequence 170587,
29	46.5	33.7	63	4	US-10-437-963-189474	Sequence 189474,
30	46.5	33.7	145	3	US-09-764-868-1233	Sequence 1233, Ap
31	46.5	33.7	444	4	US-10-104-047-2354	Sequence 2354, Ap
32	46.5	33.7	720	4	US-10-156-761-14403	Sequence 14403, A
33	46	33.3	115	6	US-10-192-381-67	Sequence 67, Appl
34	46	33.3	115	6	US-11-008-889-67	Sequence 67, Appl
35	46	33.3	140	4	US-10-424-599-280563	Sequence 280563,
36	46	33.3	149	4	US-10-425-114-46739	Sequence 46739, A
37	46	33.3	225	4	US-10-282-122A-45251	Sequence 45251, A
38	46	33.3	297	4	US-10-425-115-337842	Sequence 337842,
39	46	33.3	301	4	US-10-767-701-43907	Sequence 43907, A
40	46	33.3	310	4	US-10-223-074-11	Sequence 11, Appl
41	46	33.3	317	4	US-10-424-599-256290	Sequence 256290,
42	46	33.3	333	4	US-10-282-122A-63074	Sequence 63074, A
43	46	33.3	386	4	US-10-177-293-403	Sequence 403, Appl
44	46	33.3	393	3	US-09-823-240-11	Sequence 11, Appl
45	46	33.3	393	4	US-10-618-538-3	Sequence 3, Appli

#### ALIGNMENTS

RESULT 1  
US-10-424-599-242162  
; Sequence 242162, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 242162  
; LENGTH: 79  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_606C.1.pep  
US-10-424-599-242162

Query Match 39.9%; Score 55; DB 4; Length 79;  
Best Local Similarity 46.2%; Pred. No. 2.2;  
Matches 12; Conservative 2; Mismatches 4; Indels 8; Gaps 1;

Qy 4 RGESLNKSLP-----ILHWEKFF 21  
|||:|||||:|||||  
Db 25 RGSLNKKNPNPFLPVPVVLKEWAF 50

RESULT 2  
US-09-764-877-1036  
; Sequence 1036, Application US/09764877  
; Patent No. US20020147140A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC005  
; CURRENT APPLICATION NUMBER: US/09/764,877  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 4031  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1036  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:



```

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44667
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-10-282-122A-44667

Query Match 35.9%; Score 49.5; DB 4; Length 255;
Best Local Similarity 52.2%; Pred. No. 49;
Matches 12; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

Qy 1 NILRGESLNKSLPILHEWKFFDY 23
Db 96 NVLNRSLNTHLEALHEQQ-FDY 117

RESULT 7
US-10-418-861B-46
; Sequence 46, Application US/10418861B
; Publication No. US20040010131A1
; GENERAL INFORMATION:
; APPLICANT: da Silva, Ana Claudia Rasera
; APPLICANT: Farah, Shaker Chuck
; APPLICANT: Quaggio, Ronaldo Bento
; APPLICANT: Reinach, Fernando de Castro
; APPLICANT: Feiro, Jesus Aparecido
; APPLICANT: De Oliveira, Julio Cezar Franco
; APPLICANT: De Laia, Marcelo Luiz
; APPLICANT: Setubal Joao C.
; APPLICANT: Furian, Luiz Roberto
; TITLE OF INVENTION: Isolated Xanthomonas nucleic acid molecules, proteins encoded the
; FILE REFERENCE: uses thereof
; CURRENT APPLICATION NUMBER: US/10/418,861B
; CURRENT FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: US 60/374,620
; PRIOR FILING DATE: 2002-04-22
; NUMBER OF SEQ ID NOS: 85

```

```

; SEQ ID NO 46
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Xanthomonas
; FEATURE:
US-10-418-861B-46

Query Match 35.5%; Score 49; DB 4; Length 142;
Best Local Similarity 42.1%; Pred. No. 32;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 2 ILRGESLNKSLPILHEWKF 20
Db 80 LITGHSIRKNVPVLDAYF 98

RESULT 8
US-10-282-122A-60551
; Sequence 60551, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60551
; LENGTH: 768
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
US-10-282-122A-60551

Query Match 35.5%; Score 49; DB 4; Length 768;
Best Local Similarity 34.8%; Pred. No. 1.8e+02;
Matches 8; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

Qy 1 NILRGESLNKSLPILHEWKFFDY 23
Db 540 NLKASQSVNKLQVQYDWSADY 562

```

```
RESULT 9
US-10-437-963-106441
; Sequence 106441, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 106441
; LENGTH: 152
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(152)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_10885C.1.pep
US-10-437-963-106441
Query Match 34.8%; Score 48; DB 4; Length 152;
Best Local Similarity 53.3%; Pred. No. 1.5e+02;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 6 ESLNKSILPLHEWKF 20
Db 53 EPSKANPLLEHWHF 67

RESULT 10
US-10-424-599-190654
; Sequence 190654, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 190654
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_143180C.1.pep
US-10-424-599-190654
Query Match 34.8%; Score 48; DB 4; Length 439;
Best Local Similarity 52.6%; Pred. No. 1.5e+02;
Matches 10; Conservative 3; Mismatches 4; Indels 2; Gaps 1;

QY 5 GESLNKSLP--ILHEWKFF 21
Db 130 GESINVSLPSPSGYKWKYY 148

RESULT 11
US-10-425-114-37354
; Sequence 37354, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 37354
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: jC-gmle01810009c08_FLI.pep
US-10-425-114-37354
Query Match 34.8%; Score 48; DB 4; Length 445;
Best Local Similarity 52.6%; Pred. No. 1.5e+02;
Matches 10; Conservative 3; Mismatches 4; Indels 2; Gaps 1;

QY 5 GESLNKSLP--ILHEWKFF 21
Db 136 GESINVSLPSPSGYKWKYY 154

RESULT 12
US-10-437-963-106443
; Sequence 106443, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 106443
; LENGTH: 541
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_10887C.1.pep
US-10-437-963-106443
Query Match 34.8%; Score 48; DB 4; Length 541;
Best Local Similarity 53.3%; Pred. No. 1.8e+02;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 6 ESLNKSILPLHEWKF 20
Db 321 EPSKANPLLEHWHF 335

RESULT 13
US-10-032-585-7917
```



```
; Sequence 7917, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7917
; LENGTH: 904
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7917

Query Match      34.8%; Score 48; DB 4; Length 904;
Best Local Similarity 35.5%; Pred. No. 3.1e+02;
Matches 11; Conservative 7; Mismatches 7; Indels 6; Gaps 1;

QY      1 NILRGESLNKSL-----PILHEWKFFDYDF 25
      :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      341 SILFGLAINKTLTPAKKEVINGWKISYNP 371

RESULT 14
US-10-149-310-206
; Sequence 206, Application US/10149310
; Publication No. US20040077039A1
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas
; APPLICANT: Madden, Kevin T.
; APPLICANT: Maxon, Mary
; APPLICANT: Sherman, Amir
; TITLE OF INVENTION: Modulation of Secondary Metabolite Production by
; FILE REFERENCE: 14184-019US1
; CURRENT APPLICATION NUMBER: US/10/149,310
; CURRENT FILING DATE: 2003-02-19
; PRIOR APPLICATION NUMBER: PCT/US01/29288
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: US 60/233,564
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 206
; LENGTH: 949
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-149-310-206

Query Match      34.8%; Score 48; DB 4; Length 949;
Best Local Similarity 42.9%; Pred. No. 3.3e+02;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY      4 RGSLSNKSILPILHEWKFFDYD 24
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      369 KGDFRSQEKNIHRESKFFDFD 389

RESULT 15
US-10-425-115-201291
; Sequence 201291, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
```

```
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 201291
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_115163C.1.pep
US-10-425-115-201291

Query Match      34.1%; Score 47; DB 4; Length 46;
Best Local Similarity 43.8%; Pred. No. 20;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY      6 ESLNKSILPILHEWKFF 21
      :|:|:|:|:|:|:|:|:|:|:|:|:|
Db      31 QEFKKNLPLIQNWKLF 46

Search completed: May 3, 2006, 19:17:53
Job time : 9.77058 secs
```

This page blank (1500)

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2006, 19:12:24 ; Search time 1.44033 Seconds  
(without alignments)  
803.371 Million cell updates/sec

Title: US-10-525-567-2

Perfect score: 138

Sequence: 1 NILRGSNLKSLILHEWKFFDYDF 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 235405 seqs, 46284737 residues

Total number of hits satisfying chosen parameters: 235405

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:

- 1: /SID55/ptodata/1/pubpaa/US08 NEW PUB.pepl.\*
- 2: /SID55/ptodata/1/pubpaa/US06 NEW PUB.pep.\*
- 3: /SID55/ptodata/1/pubpaa/US07 NEW PUB.pep.\*
- 4: /SID55/ptodata/1/pubpaa/US08 NEW PUB.pep.\*
- 5: /SID55/ptodata/1/pubpaa/PCT NEW PUB.pep.\*
- 6: /SID55/ptodata/1/pubpaa/US09 NEW PUB.pep.\*
- 7: /SID55/ptodata/1/pubpaa/US09 NEW PUB.pep.\*
- 8: /SID55/ptodata/1/pubpaa/US10 NEW PUB.pep.\*
- 9: /SID55/ptodata/1/pubpaa/US10 NEW PUB.pep.\*
- 10: /SID55/ptodata/1/pubpaa/US11 NEW PUB.pep.\*
- 11: /SID55/ptodata/1/pubpaa/US11 NEW PUB.pep.\*
- 12: /SID55/ptodata/1/pubpaa/US60 NEW PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	35.5	768	11	US-11-045-004-2090
2	48	34.8	955	11	US-11-079-463-7834
3	47.5	34.4	204	11	US-11-009-658-40
4	46.5	33.7	444	11	US-11-072-512-2354
5	46	33.3	238	11	US-11-096-568A-33583
6	46	33.3	268	11	US-11-079-463-8150
7	46	33.3	314	11	US-11-096-568A-33582
8	46	33.3	334	11	US-11-096-568A-33581
9	45	32.6	135	11	US-11-096-568A-1311
10	45	32.6	538	11	US-11-087-099-1023
11	45	32.6	538	11	US-11-087-099-8465
12	45	32.6	538	11	US-11-087-099-11979
13	45	32.6	538	11	US-11-188-298-18858
14	44	31.9	201	11	US-11-188-298-18858
15	44	31.9	219	11	US-11-087-099-6165
16	44	31.9	240	11	US-11-179-977-19
17	44	31.9	260	11	US-11-096-568A-6893
18	44	31.9	305	9	US-10-063-703-108
19	44	31.9	305	9	US-10-194-487-324
20	44	31.9	305	9	US-10-195-883-324
21	44	31.9	305	9	US-10-195-888-324

Sequence 324, App  
Sequence 108, App  
Sequence 108, App  
Sequence 4, Appli  
Sequence 2, Appli  
Sequence 8270, Ap  
Sequence 17331, A  
Sequence 1472, Ap  
Sequence 10168, A  
Sequence 102, App  
Sequence 1745, Ap  
Sequence 1746, Ap  
Sequence 6, Appli  
Sequence 7687, Ap  
Sequence 1220, Ap  
Sequence 5736, Ap  
Sequence 8020, Ap  
Sequence 12109, A  
Sequence 11154, A  
Sequence 12172, A  
Sequence 16274, A  
Sequence 18448, A  
Sequence 4, Appli

# ALIGNMENTS

## RESULT 1

US-11-045-004-2090  
Sequence 2090, Application US/11045004  
Publication No. US20060078901A1

GENERAL INFORMATION:

- APPLICANT: BUCHRIESEN, CARMEN
- APPLICANT: FRANGEUL, LIONEL
- APPLICANT: COUVE, ELISABETH
- APPLICANT: RUSNIOK, CHRISTOPHE
- APPLICANT: FSIHI, HAFIDA
- APPLICANT: DEHOUX, PIERRE
- APPLICANT: DUSSURGET, OLIVIER
- APPLICANT: CHETOUANI, FARID
- APPLICANT: NEDJARI, HAFED
- APPLICANT: GLASER, PHILIPPE
- APPLICANT: KUNST, FRANCK
- APPLICANT: COSSART, PASCALE
- APPLICANT: DANIELS, JUSTIN
- APPLICANT: GOEBEL, WERNER
- APPLICANT: KREFT, JURGEN
- APPLICANT: KUHN, MICHAEL
- APPLICANT: NG, EVA
- APPLICANT: VAZQUEZ-BOLAND, ANTONIO
- APPLICANT: DOMINGUEZ-BERNAL, GUSTAVO
- APPLICANT: GARRIDO-GARCIA, PATRICIA
- APPLICANT: TIERREZ-MARTINEZ, ALBERTO
- APPLICANT: AMEND, ALEXANDRA
- APPLICANT: CHAKRABORTY, TRINAD
- APPLICANT: DOMANN, EUGEN
- APPLICANT: HAIN, THORSTEN
- APPLICANT: BERCHE, PATRICK
- APPLICANT: CHARBIT, ALAIN
- APPLICANT: DURANT, LIONEL
- APPLICANT: PEREZ-DIAZ, JOSE-CLAUDIO
- APPLICANT: BAQUERO, FERNANDO
- APPLICANT: GARCIA DEL PORTILLO, FRANCISCO
- APPLICANT: GOMEZ-LOPEZ, NURIA
- APPLICANT: MADUENIO, ENCARNNA
- APPLICANT: PABLOS, BETRIZ DE
- APPLICANT: WEHLAND, JURGEN
- APPLICANT: KARST, UWE
- APPLICANT: ENTIAN, KARL-DIETER
- APPLICANT: HAUF, JORG
- APPLICANT: ROSE, MATTHIAS

```

; APPLICANT: VOSS, HAMUT
; TITLE OF INVENTION: LISTERIA MONOCYTOGENES GENOME, POLYPEPTIDES AND USES
; FILE REFERENCE: 05394.0018-02
; CURRENT APPLICATION NUMBER: US/11/045,004
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: 10/637,657
; PRIOR FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: 10/257,023
; PRIOR FILING DATE: 2002-10-08
; PRIOR APPLICATION NUMBER: PCT/FR01/01118
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: FR 00/04,629
; PRIOR FILING DATE: 2000-04-11
; NUMBER OF SEQ ID NOS: 2854
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2090
; LENGTH: 768
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
US-11-045-004-2090

Query Match      35.5%; Score 49; DB 11; Length 768;
Best Local Similarity 34.8%; Pred. No. 13;
Matches 8; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 NLRGSLNKSILPILHEWKFFDY 23
   : : : : : : : : : : : : : : : :
Db 540 NLRASQSVNKLQVQVDSSADY 562

RESULT 2
US-11-079-463-7834
; Sequence 7834, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR
; FILE REFERENCE: PATH00-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 7834
; LENGTH: 955
; TYPE: PRT
; ORGANISM: B.fragilis
US-11-079-463-7834

Query Match      34.8%; Score 48; DB 11; Length 955;
Best Local Similarity 53.3%; Pred. No. 24;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 4 RGSLSNKSILPILHEW 18
   : : : : : : : : : : : : : : :
Db 138 KDEALDRSLILHDW 152

RESULT 3
US-11-009-658-40
; Sequence 40, Application US/11009658
; Publication No. US20060003430A1
; GENERAL INFORMATION:
; APPLICANT: E. I. du Pont de Nemours and Company
; TITLE OF INVENTION: Lysophosphatidic Acid Acetyltransferases
; FILE REFERENCE: BB1332
; CURRENT APPLICATION NUMBER: US/11/009,658
; CURRENT FILING DATE: 2004-12-10
; PRIOR APPLICATION NUMBER: US/09/914,098
; PRIOR FILING DATE: 2001-08-22

```

```

; PRIOR APPLICATION NUMBER: 60/121,119
; PRIOR FILING DATE: 1999-02-22
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 40
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (68)
; OTHER INFORMATION: Xaa = any amino acid
US-11-009-658-40

Query Match      34.4%; Score 47.5; DB 11; Length 204;
Best Local Similarity 47.1%; Pred. No. 5.1;
Matches 8; Conservative 5; Mismatches 3; Indels 1; Gaps 1;

QY 7 SLNKSILPILHEWKFFDY 23
   : : : : : : : : : : : : : : :
Db 168 ALN-ALPTYYSWRFYDY 183

RESULT 4
US-11-072-512-2354
; Sequence 2354, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, KYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTYUUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2354
; LENGTH: 444
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-2354

Query Match      33.7%; Score 46.5; DB 11; Length 444;
Best Local Similarity 30.0%; Pred. No. 18;
Matches 12; Conservative 1; Mismatches 10; Indels 17; Gaps 1;

QY 3 LRGESL-----NKSILPILHEWKFFDYDF 25
   : : : : : : : : : : : : : : :
Db 60 IRGERLRDKWVLECMLEKDLIYNKVTPTFHHWKIDDKKF 99

RESULT 5
US-11-096-568A-33583

```

```
; Sequence 33583, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 33583
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(238)
; OTHER INFORMATION: Ceres Seq. ID no. 13603494
US-11-096-568A-33583

Query Match      33.3%; Score 46; DB 11; Length 238;
Best Local Similarity 43.8%; Pred. No. 11;
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy      8 LNKSLPILHEWKFFDY 23
      : : |||
Db     34 IQRFLEPQQSWTFDY 49

RESULT 6
US-11-079-463-8150
; Sequence 8150, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FR
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PATH00-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 8150
; LENGTH: 268
; TYPE: PRT
; ORGANISM: B.fragilis
US-11-079-463-8150

Query Match      33.3%; Score 46; DB 11; Length 268;
Best Local Similarity 38.9%; Pred. No. 12;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy      8 LNKSLPILHEWKFFDYDF 25
      : : ||| : |||
Db     186 LNRHTGKIAEWQYDNY 203

RESULT 7
US-11-096-568A-33582
; Sequence 33582, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 33582
```

```
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(314)
; OTHER INFORMATION: Ceres Seq. ID no. 13603493
US-11-096-568A-33582

Query Match      33.3%; Score 46; DB 11; Length 314;
Best Local Similarity 43.8%; Pred. No. 14;
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy      8 LNKSLPILHEWKFFDY 23
      : : |||
Db     110 IQRFLEPQQSWTFDY 125

RESULT 8
US-11-096-568A-33581
; Sequence 33581, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 33581
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(334)
; OTHER INFORMATION: Ceres Seq. ID no. 13603492
US-11-096-568A-33581

Query Match      33.3%; Score 46; DB 11; Length 334;
Best Local Similarity 43.8%; Pred. No. 15;
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy      8 LNKSLPILHEWKFFDY 23
      : : |||
Db     130 IQRFLEPQQSWTFDY 145

RESULT 9
US-11-096-568A-1311
; Sequence 1311, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 1311
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(135)
; OTHER INFORMATION: Ceres Seq. ID no. 14302141
US-11-096-568A-1311

Query Match      32.6%; Score 45; DB 11; Length 135;
Best Local Similarity 46.7%; Pred. No. 8.1;
```

```
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 GESLNKSLPILHEWK 19
   | : | : | : |
Db 28 GITLRQHPIPIFHWK 42

RESULT 10
US-11-087-099-1023
; Sequence 1023, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 1023
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Candida maltosa
US-11-087-099-1023

Query Match 32.6%; Score 45; DB 11; Length 538;
Best Local Similarity 45.5%; Pred. No. 37;
Matches 10; Conservative 3; Mismatches 7; Indels 2; Gaps 1;

QY 2 ILRGESLNKSLPILHEWKFFDY 23
   ||| ||| : ||| :
Db 270 ILNGSKFNKSIKTVH--KPADF 289

RESULT 11
US-11-087-099-8465
; Sequence 8465, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 8465
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Candida maltosa
US-11-087-099-8465

Query Match 32.6%; Score 45; DB 11; Length 538;
Best Local Similarity 45.5%; Pred. No. 37;
Matches 10; Conservative 3; Mismatches 7; Indels 2; Gaps 1;

QY 2 ILRGESLNKSLPILHEWKFFDY 23
   ||| ||| : ||| :
Db 270 ILNGSKFNKSIKTVH--KPADF 289

RESULT 12
US-11-188-298-11979
; Sequence 11979, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 11979
```

```
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Candida maltosa
US-11-188-298-11979

Query Match 32.6%; Score 45; DB 11; Length 538;
Best Local Similarity 45.5%; Pred. No. 37;
Matches 10; Conservative 3; Mismatches 7; Indels 2; Gaps 1;

QY 2 ILRGESLNKSLPILHEWKFFDY 23
   ||| ||| : ||| :
Db 270 ILNGSKFNKSIKTVH--KPADF 289

RESULT 13
US-11-188-298-18858
; Sequence 18858, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 18858
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Candida maltosa
US-11-188-298-18858

Query Match 32.6%; Score 45; DB 11; Length 538;
Best Local Similarity 45.5%; Pred. No. 37;
Matches 10; Conservative 3; Mismatches 7; Indels 2; Gaps 1;

QY 2 ILRGESLNKSLPILHEWKFFDY 23
   ||| ||| : ||| :
Db 270 ILNGSKFNKSIKTVH--KPADF 289

RESULT 14
US-11-096-568A-6894
; Sequence 6894, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 6894
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(201)
; OTHER INFORMATION: Ceres Seq. ID no. 15168590
US-11-096-568A-6894

Query Match 31.9%; Score 44; DB 11; Length 201;
Best Local Similarity 39.1%; Pred. No. 18;
Matches 9; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

QY 3 LRGESLNKSLPILHEWKFFDYDF 25
   ||| ||| : ||| :
Db 146 LRGFDLNLPAPLTFEWSMGDF 168
```

```
RESULT 15
US-11-087-099-6165
; Sequence 6165, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 6165
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Nicotiana plumbaginifolia
US-11-087-099-6165

Query Match      31.9%; Score 44; DB 11; Length 219;
Best Local Similarity 39.1%; Pred. No. 20;
Matches 9; Conservative 5; Mismatches 7; Indels 2; Gaps 1;

Qy      1 NILGESLNKSLPILHEWKPFYD 23
Db      188 NCMENERVSKSLP--HPHKIYDP 208

Search completed: May 3, 2006, 19:18:56
Job time : 3.44033 secs
```

This Page Blank (user)



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2006, 18:53:48 ; Search time 222.646 Seconds  
(without alignments)  
1034.083 Million cell updates/sec

Title: US-10-525-567-3  
Perfect score: 2796  
Sequence: 1 AAVNQRKSNLAHSMKVI.....NNQNDNRNDNQVHSSKLH 524

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_21.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*  
9: Geneseqp2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2796	100.0	524	8	ADL27606 Honey bee
2	2796	100.0	544	8	ADL27616 Honey bee
3	2773	99.2	544	8	ADL27618 Honey bee
4	1420.5	50.8	414	8	ADL27617 Honey bee
5	1415.5	50.6	413	8	ADL27607 Honey bee
6	1415.5	50.6	432	4	AAQ78710 Oxygen pr
7	1415.5	50.6	432	4	AAQ78815 Bee prote
8	1415.5	50.6	432	8	ADL27619 Honey bee
9	427	15.3	541	4	ABB60127 Drosophil
10	422.5	15.1	426	4	ABB66147 Drosophil
11	418.5	15.0	453	4	ABB71124 Drosophil
12	415.5	14.9	412	4	ABB64733 Drosophil
13	408.5	14.6	415	4	ABB64727 Drosophil
14	408.5	14.6	438	4	ABB60071 Drosophil
15	322.5	11.5	256	2	AAQ24796 Sequence
16	320	11.4	429	4	ABB67082 Drosophil
17	319.5	11.4	409	4	ABB66148 Drosophil
18	313.5	11.2	342	2	AAW03625 Human lut
19	302.5	10.8	379	4	ABB71640 Drosophil
20	294.5	10.5	316	4	ABB58294 Drosophil
21	290	10.4	590	2	AAK11019 Apo-lipop
22	289.5	10.4	293	9	AEB40378 L. pneumo
23	282.5	10.1	319	9	AEB37025 L. pneumo
24	281	10.1	530	4	ABB71802 Drosophil

25	267	9.5	2150	5	AAO22566	Aao22566 Wooden le
26	261	9.3	190	2	AAK41992	Aar41992 Staphyloc
27	261	9.3	1712	3	AAK18205	Aab18205 Plasmodiu
28	260	9.3	190	2	AAK42004	Aar42004 Staphyloc
29	258	9.2	190	2	AAK41996	Aar41996 Staphyloc
30	258	9.2	190	2	AAK42012	Aar42012 Staphyloc
31	258	9.2	190	2	AAK41995	Aar41995 Staphyloc
32	257	9.2	190	2	AAK42011	Aar42011 Staphyloc
33	257	9.2	190	2	AAK41994	Aar41994 Staphyloc
34	257	9.2	190	2	AAK42010	Aar42010 Staphyloc
35	256	9.2	190	2	AAK41998	Aar41998 Staphyloc
36	256	9.2	190	2	AAK42002	Aar42002 Staphyloc
37	256	9.2	190	2	AAK41989	Aar41989 Staphyloc
38	256	9.2	190	2	AAK41993	Aar41993 Staphyloc
39	256	9.2	190	2	AAK41991	Aar41991 Staphyloc
40	256	9.2	190	2	AAK41997	Aar41997 Staphyloc
41	256	9.2	190	2	AAK42003	Aar42003 Staphyloc
42	255	9.1	190	2	AAK41990	Aar41990 Staphyloc
43	255	9.1	190	2	AAK41999	Aar41999 Staphyloc
44	255	9.1	190	2	AAK42013	Aar42013 Staphyloc
45	255	9.1	190	2	AAK42005	Aar42005 Staphyloc

ALIGNMENTS

RESULT 1  
ADL27606  
ID ADL27606 standard; protein; 524 AA.  
XX  
AC ADL27606;  
XX  
XX  
DT 03-JUN-2004 (first entry)  
XX  
DE Honey bee RJP70 mature protein, SEQ ID 3.  
XX  
KW Antiallergic; Antiasthmatic; Antiinflammatory; Dermatological; honey bee;  
KW royal jelly protein; allergy; pollinosis; atopic dermatitis;  
KW contact hypersensitivity; bronchial asthma; allergic rhinitis; RJP70.  
XX  
OS Apis mellifera.  
XX  
PN WO2004019971-A1.  
XX  
PD 11-MAR-2004.  
XX  
PF 26-AUG-2003; 2003WO-JP010795.  
XX  
PR 29-AUG-2002; 2002JP-00252087.  
XX  
PR 30-JAN-2003; 2003JP-00022776.  
XX  
(HAYA/) HAYASHIBARA K.  
XX  
OKamoto I, Arai N, Kohno K, Kurimoto M, Sano O;  
XX  
WPI; 2004-248191/23.  
XX  
N-ESDB; ADL27608.  
XX  
Antiallergic agents for relieving symptoms accompanying e.g. pollinosis,  
XX atopic dermatitis, contact hypersensitivity, bronchial asthma and  
XX allergic rhinitis, comprise protein originated from royal jelly.  
XX  
Claim 1; SEQ ID NO 3; 78pp; Japanese.  
XX  
The present invention relates to novel antiallergic agents, which  
XX comprise as an active ingredient, peptides ADL27604 and ADL27605, which  
XX are N-terminal peptides of honey bee royal jelly proteins (RJP, ADL27606  
XX and ADL27607). The agents can be used to relieve symptoms accompanying an  
XX allergic disease e.g. pollinosis, atopic dermatitis, contact  
XX hypersensitivity, bronchial asthma and allergic rhinitis, which are  
XX applicable in foods, cosmetics and drugs.  
XX  
Sequence 524 AA;  
SQ



DE Honey bee MRJP3.

XX Antiallergic; Antiasthmatic; Antiinflammatory; Dermatological; honey bee;

KW royal jelly protein; allergy; pollinosis; atopic dermatitis;

KW contact hypersensitivity; bronchial asthma; allergic rhinitis; MRJP3.

XX

XX Apis mellifera.

OS

XX

FH Key Location/Qualifiers

FT Peptide 1..16

FT Protein /label= Signal\_peptide

FT 17..544

FT /label= Mature\_protein

XX

PN WO2004019971-A1.

XX

XX 11-MAR-2004.

XX

XX 26-AUG-2003; 2003WO-JP010795.

XX

XX 29-AUG-2002; 2002JP-00252087.

PR 30-JAN-2003; 2003JP-00022776.

PR

XX (HAYA/) HAYASHIBARA K.

PA

XX Okamoto I, Arai N, Kohno K, Kurimoto M, Sano O;

PI

XX WPI; 2004-248191/23.

XX N-PSDB; ADL27610.

DR

XX Antiallergic agents for relieving symptoms accompanying e.g. pollinosis,

XX atopic dermatitis, contact hypersensitivity, bronchial asthma and

PT allergic rhinitis, comprise protein originated from royal jelly.

PT

XX

PS Disclosure; Page 65-69; 78pp; Japanese.

PS

XX

XX The present invention relates to novel antiallergic agents, which

CC comprise as an active ingredient, peptides ADL27604 and ADL27605, which

CC are N-terminal peptides of honey bee royal jelly proteins (RJP, ADL27606

CC and ADL27607). The agents can be used to relieve symptoms accompanying an

CC allergic disease e.g. pollinosis, atopic dermatitis, contact

CC hypersensitivity, bronchial asthma and allergic rhinitis, which are

CC applicable in foods, cosmetics and drugs. The present sequence is a honey

CC bee major RJP.

XX

XX Sequence 544 AA;

SQ

Query Match 99.2%; Score 2773; DB 8; Length 544;

Best Local Similarity 99.4%; Pred. No. 8e-218;

Matches 521; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAVNHQKSAANLAHSMKVIYEWKHIDPFGSDERRDAAIKSGEPDHTKYPFDVDRWRD 60

DB 21 AAVNHQKSAANLAHSMKVIYEWKHIDPFGSDERRDAAIKSGEPDHTKYPFDVDRWRD 80

QY 61 KTFVTIERNNGVPSLNVNKKGGKGLLRPPDWSFAKYEDCSGIYSAFKIADVDEDR 120

DB 81 KTFVTIERNNGVPSLNVNKKGGKGLLRPPDWSFAKYEDCSGIYSAFKIADVDEDR 140

QY 121 LWVLDSGLVNNQPCSKLLTDFDLTKSLVKQVEIPHNIAVNTTGMGELVSLVAQVAD 180

DB 141 LWVLDSGLVNNQPCSKLLTDFDLTKSLVKQVEIPHNIAVNTTGMGELVSLVAQVAD 200

QY 181 RNTNVTYIADKEGELIMYQNSDDSFHLRTSTNTFDYDPRYTKLTIVAGESFTVKNGICGIA 240

DB 201 RNTNVTYIADKEGELIMYQNSDDSFHLRTSTNTFDYDPRYTKLTIVAGESFTVKNGIYGIA 260

QY 241 LSPVTNNLYSPSLSHGLYYVDTQFRNPQYENNVQVEGSDIINTQSGFKVWSKNGVL 300

DB 261 LSPVTNNLYSPSLSHGLYYVDTQFRNPQYENNVQVEGSDIINTQSGFKVWSKNGVL 320

QY 301 FLGLVNGSGIACVNEHQVLORESFDVVAQNEETLQWIVSMKIMENLPQSGRIINDEPGEY 360

321 FLGLVNGSGIACVNEHQVLORESFDVVAQNEETLQWIVSMKIMENLPQSGRIINDEPGEY 380

361 MLALSNRMOKIINNDFNFNDVNFRIILGANVDDLNRNTRCGRYHQNAGNQADNQADNQ 420

381 MLALSNRMOKIINNDFNFNDVNFRIILGANVDDLNRNTRCGRYHQNAGNQADNQADNQ 440

421 NANNQNADNQANKQNGNRQNDNRQNDNRQNDNRQNDNRQNDNRQNDNRQNDNRQNGK 480

441 NANNQNADNQANKQNGNRQNDNRQNDNRQNDNRQNDNRQNDNRQNDNRQNDNRQNGK 500

481 NDNKQNGNRQNDNRQNGNRQNDNRQNDNRQNDNRQNDNRQNDNRQNDNRQNDNRQNGK 524

501 NDNKQNGNRQNDNRQNGNRQNDNRQNDNRQNDNRQNDNRQNDNRQNDNRQNDNRQNGK 544

RESULT 4

ADL27617

ID ADL27617 standard; protein; 414 AA.

XX

AC ADL27617;

XX

DT 03-JUN-2004 (first entry)

XX

DE Honey bee RJP55.

XX

XX Antiallergic; Antiasthmatic; Antiinflammatory; Dermatological; honey bee;

KW royal jelly protein; allergy; pollinosis; atopic dermatitis;

KW contact hypersensitivity; bronchial asthma; allergic rhinitis; RJP55.

XX

OS Apis mellifera.

PN WO2004019971-A1.

XX

XX 11-MAR-2004.

XX

XX 26-AUG-2003; 2003WO-JP010795.

XX

XX 29-AUG-2002; 2002JP-00252087.

PR 30-JAN-2003; 2003JP-00022776.

XX (HAYA/) HAYASHIBARA K.

PA

XX Okamoto I, Arai N, Kohno K, Kurimoto M, Sano O;

PI

XX WPI; 2004-248191/23.

XX N-PSDB; ADL27609.

DR

XX Antiallergic agents for relieving symptoms accompanying e.g. pollinosis,

XX atopic dermatitis, contact hypersensitivity, bronchial asthma and

PT allergic rhinitis, comprise protein originated from royal jelly.

PT

XX

PS Disclosure; Page 62-65; 78pp; Japanese.

PS

XX

XX The present invention relates to novel antiallergic agents, which

CC comprise as an active ingredient, peptides ADL27604 and ADL27605, which

CC are N-terminal peptides of honey bee royal jelly proteins (RJP, ADL27606

CC and ADL27607). The agents can be used to relieve symptoms accompanying an

CC allergic disease e.g. pollinosis, atopic dermatitis, contact

CC hypersensitivity, bronchial asthma and allergic rhinitis, which are

CC applicable in foods, cosmetics and drugs. The present sequence is a honey

CC bee RJP.

XX

XX Sequence 414 AA;

SQ

Query Match 50.8%; Score 1420.5; DB 8; Length 414;

Best Local Similarity 65.3%; Pred. No. 2.3e-107;

Matches 265; Conservative 63; Mismatches 65; Indels 13; Gaps 2;

QY 12 NLAHSMKVIYEWKHIDPFGSDERRDAAIKSGEPDHTKYPFDVDRWRDFTVTIERNNG 71

DB 8 SLNKSPLTILHEWKFFDFGSDERRDAAIKSGEPDHTKYPFDVDRWRDFTVTIERNNG 67

QY 72 VPSSILNVNKKGGKGLLRPPDWSFAKYEDCSGIVSAFKIAVDKFDRLWVLDSGLVNN 131



PS	Disclosure; Page 9-10; 10pp; Japanese.
XX	The present invention relates to a lactic acid accumulation inhibitor.
CC	The inhibitor comprises a protein contained in royal jelly having lactic
CC	acid accumulation inhibiting activity as the active component. The lactic
CC	acid accumulation inhibitor is used for restoration, improvement and
CC	release of muscular fatigue. The present sequence is a bee protein, which
CC	was used in the present invention
XX	Sequence 432 AA;
QY	Query Match 50.6%; Score 1415.5; DB 4; Length 432;
DB	Best Local Similarity 65.0%; Pred. No. 6.3e-107; Indels 13; Gaps 2;
DB	Matches 264; Conservative 63; Mismatches 66; Indels 13; Gaps 2;
QY	12 NLAHSMKVIYEWKHIDPFGSDERRDAAIKSGEFDHTKNTYPPFDVDRWRDKTFVTIERNNG 71
DB	26 SLNKSILPILHEWKFFDYDFGSDERRDQDAILSGEYDYKNNYPSDIDQWHDKIFVTMLRYNG 85
QY	72 VPSSLNVVTKKGGPGLLRYPDPWSFAKYEDCSGIVSAFKIAVDKFDRLWLDLGLVNN 131
DB	86 VPSSLNVISKVGDGGLLPQYPDPWSFAKYDDCSGIVSASKLAIDKCDRLWLDLGLVNN 145
QY	132 NOPCMSPKLLTFDLTKTSKLVKQVEIPHNTAVNATTGMLVSLAVQAID--RTNTMVI 188
DB	146 TQPMCSPKLLTFDLTKTSKLVKQVEIPHNVAVNATTGKRLSSLAQSLDCNTNSDTMVI 205
QY	189 ADEKSGELIMYQNSDDSFHRLTSNTFDYDPRTYKLTVAGSEFTVKGICGIALSPVTNNL 248
DB	206 ADEKSGELIVYHNSDDSFHRLTSNTFDYDPKFTKMTIDGESYTAQDGISGVALSPMTNNL 265
QY	249 YYSPLSHGLIYVDTEQFRNPQYENNVQYEGSQDILNTQSGFKVSKNGVLPFLGLVGN 308
DB	266 YYSFVASTSLYYVNTQFRTSDYQQNDIHVEGVQNILDTQSSAKVSKSGVLPFLGLVGD 325
QY	309 GIACVNEHQVLORESFDVVAQNEETLQIMVSKMIMENLPQSGRINDPEGNEYMLALSIRM 368
DB	326 ALGCWNEHRTLERHNTVARTVAQSDTELQMIASMKIKEAXPHVPIDRYINREYILVLSNKM 385
QY	369 QKIINDDFNDVNFRIIGANVDDLMRNTRCGRYHQNAGNQADN 414
DB	386 QXWVNDNFDDVNFRIIMANVNELILNTRC-----ENPDN 421
RESULT 8	
ADL27619	
ID	ADL27619 standard; protein; 432 AA.
AC	ADL27619;
XX	03-JUN-2004 (first entry)
DE	Honey bee MRJPl.
XX	Antiallergic; Antiasthmatic; Antiinflammatory; Dermatological; honey bee;
KW	royal jelly protein; allergy; pollinosis; atopic dermatitis;
KW	contact hypersensitivity; bronchial asthma; allergic rhinitis; MRJPl.
XX	Apis mellifera.
OS	Key Location/Qualifiers
FT	Peptide 1..19
FT	/label= Signal_peptide
FT	Protein 20..432
FT	/label= Mature_protein
XX	WO2004019971-A1.
XX	11-MAR-2004.
XX	26-AUG-2003; 2003WO-JP010795.
XX	29-AUG-2002; 2002JP-00252087.

Query Match	50.6%; Score 1415.5; DB 4; Length 432;
Best Local Similarity	65.0%; Pred. No. 6.3e-107;
Matches	264; Conservative 63; Mismatches 66; Indels 13; Gaps 2;
QY	12 NLAHSMKVIYEWKHIDPFGSDERRDAAIKSGEFDHTKNTYPPFDVDRWRDKTFVTIERNNG 71
DB	26 SLNKSILPILHEWKFFDYDFGSDERRDQDAILSGEYDYKNNYPSDIDQWHDKIFVTMLRYNG 85
QY	72 VPSSLNVVTKKGGPGLLRYPDPWSFAKYEDCSGIVSAFKIAVDKFDRLWLDLGLVNN 131
DB	86 VPSSLNVISKVGDGGLLPQYPDPWSFAKYDDCSGIVSASKLAIDKCDRLWLDLGLVNN 145
QY	132 NOPCMSPKLLTFDLTKTSKLVKQVEIPHNTAVNATTGMLVSLAVQAID--RTNTMVI 188
DB	146 TQPMCSPKLLTFDLTKTSKLVKQVEIPHNVAVNATTGKRLSSLAQSLDCNTNSDTMVI 205
QY	189 ADEKSGELIMYQNSDDSFHRLTSNTFDYDPRTYKLTVAGSEFTVKGICGIALSPVTNNL 248
DB	206 ADEKSGELIVYHNSDDSFHRLTSNTFDYDPKFTKMTIDGESYTAQDGISGVALSPMTNNL 265
QY	249 YYSPLSHGLIYVDTEQFRNPQYENNVQYEGSQDILNTQSGFKVSKNGVLPFLGLVGN 308
DB	266 YYSFVASTSLYYVNTQFRTSDYQQNDIHVEGVQNILDTQSSAKVSKSGVLPFLGLVGD 325
QY	309 GIACVNEHQVLORESFDVVAQNEETLQIMVSKMIMENLPQSGRINDPEGNEYMLALSIRM 368
DB	326 ALGCWNEHRTLERHNTVARTVAQSDTELQMIASMKIKEAXPHVPIDRYINREYILVLSNKM 385
QY	369 QKIINDDFNDVNFRIIGANVDDLMRNTRCGRYHQNAGNQADN 414
DB	386 QXWVNDNFDDVNFRIIMANVNELILNTRC-----ENPDN 421
RESULT 7	
AAG78815	
ID	AAG78815 standard; protein; 432 AA.
AC	AAG78815;
XX	27-NOV-2001 (first entry)
DE	Bee protein.
XX	Bee; lactic acid accumulation inhibitor; royal jelly; muscular fatigue.
OS	Apis mellifera.
XX	Key Location/Qualifiers
FT	Peptide 1..19
FT	/label= Signal_peptide
FT	Protein 20..432
FT	/label= Mature_bee_protein
FT	Misc-difference 363
FT	/label= unknown
FT	/note= "Encoded by CTN, given as Xaa in specification"
XX	JP2001172190-A.
XX	26-JUN-2001.
XX	15-DEC-1999; 99JP-00356006.
XX	15-DEC-1999; 99JP-00356006.
XX	(POKK ) POLA CHEM IND INC.
XX	WPI; 2001-592556/67.
DR	N-PSDB; AAI65041.
XX	Lactic acid accumulation inhibitor comprises protein contained in royal
PT	jelly, useful for restoration, improvement and release of muscular
PT	fatigue.
XX	



```
AC ABB66147;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 25233.
XX
DE Drosophila; developmental biology; cell signalling; insecticide;
XX
DE Drosophila melanogaster.
XX
DE WO200171042-A2.
XX
DE 27-SEP-2001.
XX
DE 23-MAR-2001; 2001WO-US009231.
XX
DE 23-MAR-2000; 2000US-0191637P.
XX
DE 11-JUL-2000; 2000US-00614150.
XX
DE (PEKE ) PE CORP NY.
XX
DE Venter JC, Adams M, Li PWD, Myers EW;
XX
DE WPI; 2001-656860/75.
XX
DE N-PSDB; ABL10250.
XX
DE New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions.
XX
DE Disclosure; SEQ ID NO 25233; 2lpp + Sequence Listing; English.
XX
PS
SS
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
CC ABBS72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 426 AA;
Query Match 15.1%; Score 422.5; DB 4; Length 426;
Best Local Similarity 26.2%; Pred. No. 1e-25;
Matches 111; Conservative 83; Mismatches 155; Indels 75; Gaps 13;
QY 19 VIYWKHIDPFGSDERRDAAIKSGEFHTKXNYPFVDWRW-----RDKTFTVIER-NNG 71
DB 34 IVFEMKNLQYGFPSQERDQVLRNGRYNPDSPFIPIDIVYPPNGGPPRHFTVTSRFGQG 93
QY 72 VPSSLNVNKKGGGGLLRYP--DWSFAKEDCSGIVSAFKAVDKFRLWLDSGLV 129
DB 94 VPFSLGYTVNQRENGSEIOAYPSYQWSSHGANGDGLTSYVRHIDACGMWVLDSEI 153
QY 130 NNNQPMCSKLLTFLDKTSKLVKQVEIPHNIATVAVNATTGMGLVSLAVQAIDR-----TN 183
DB 154 EFVQ-HCAPOVMVFLDQLIHRYLPE---TSYKAKVSFVNIADIRPPSGQCKD 209
QY 184 TMVYIADKRGGLIYQNSDPSFRLTNTFTDY-DPRYTKLTVAGESFTVKNIGCIGALS 242
DB 210 VFAYLADTSAIVYDVVGQSSWRI-ENKFTYPAKFGTHTVAGESFELLDGPLALATT 268
QY 243 P-----VTNNLYSPUSLHGLYVDTEQRPNQYENNVQYEGSQDILNT----- 287
DB 269 PLGLGLRRHLIFHALS-----NELELAIPDLILNATNWKQGLSSS 309
QY 288 -----QSGKGVSKNGVFLGLVGNNGSIACVNEHQVLORESFDVVVAQNEETLQ 335
DB 310 LSEFTVLGKRGICQASHAISROGFLFCGFLPEIGFGWDIRRPNYNNVKLAINPATLQ 369
```

```
QY 336 MIVSMKIMENLPQSGRINDPEGNEVYMLALSNRMQKIINDFNFDVNFRIILGANVDDLMR 395
DB 370 FVSGMKIVRR-PADGR-----EELWLSDRLOKIFAGTIDYREINRYVMRCVDDLLQ 421
QY 396 NTRC 399
DB 422 GRGC 425
RESULT 11
ID ABB71124 standard; protein; 453 AA.
XX
AC ABB71124;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 40164.
XX
DE Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX
DE Drosophila melanogaster.
XX
DE WO200171042-A2.
XX
DE 27-SEP-2001.
XX
DE 23-MAR-2001; 2001WO-US009231.
XX
DE 23-MAR-2000; 2000US-0191637P.
XX
DE 11-JUL-2000; 2000US-00614150.
XX
DE (PEKE ) PE CORP NY.
XX
DE Venter JC, Adams M, Li PWD, Myers EW;
XX
DE WPI; 2001-656860/75.
XX
DE N-PSDB; ABL15227.
XX
DE New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions.
XX
DE Disclosure; SEQ ID NO 40164; 2lpp + Sequence Listing; English.
XX
PS
SS
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
CC ABBS72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 453 AA;
Query Match 15.0%; Score 418.5; DB 4; Length 453;
Best Local Similarity 27.9%; Pred. No. 2.4e-25;
Matches 122; Conservative 84; Mismatches 194; Indels 37; Gaps 12;
QY 16 SMKVIYWKHIDPFGSDERRDAAIKSGEFHTKXNYPFVDWRWDRDKTFTVIER-NNGVPS 74
DB 21 NLRVAYEWREMDFKYANPDQKWSAIERGEFKPANVIPFGLVAGHRLFTVLPWRDGVPA 80
QY 75 SLNVV-TNKKGGGGLLRYPYPDWSFAKYEDCS-CIVSAFKIADVDFRLWLD---SGLV 129
DB 81 SLAYLDLNDTSSKGPALKPPFSWQAHNLQEAPELVSFPRVADRRCGLRWLDSRISGLV 140
QY 130 NNNQPMCSKLLTFLDKTSKLVKQVEIPHNIATVAVNATTGMGLVSLAVQAIDRTNTMYIA 189
```



```
Db 141 EQTKIYGAQLLYVDLHNDLLRRHVLPAQOLKQGSLL----LANLAVEDSDCENTFAYAA 196
Qy 190 DERKEGLIMTQNSDDSFHRLTSNTFDYDPRYTKLTAVAGESFTVKNGICGIALS-PVTN-- 246
Db 197 DLGSPGLVYVSWKDEESWRVQHFFHDPDPWAGNFSINGIEFQWDDGLGLALSKELETGY 256
Qy 247 -NLYSPLSHGLYVYDTEQFRNPQVEENNVOYE-----GSQDILNTQSGKVVSKN-GV 299
Db 257 ATLYFHLPLCSTTFBSVDTSLRNKTLATSPMIYREFKVLGSRG-PNTQAGAEFLDPTGV 315
Qy 300 LFLGLVNSGIACVNEHVLQRESFDVVAQBEETLQMVSMKIMENLPQSGRINDEPGE 359
Db 316 LFVALPMLNEVACWRTATDSHSQSQRHNNDTLVFPSPDIKV-----DDQK 362
Qy 360 YMLALSNRMKIINNDFNFVNFRIILGANVDDLMRNTCRGCRHYHQNQAGNQADNQVADN 419
Db 363 RLWVLSNQLPVFIYDELYAGSINFRILTASVKEAIENTAC---EIRTSPLPDVINKLGD 419
Qy 420 QNANNQNADNONANKON 436
Db 420 LNTNIKLKSNSASSLRN 436

RESULT 12
ABBE64733
ID ABBE64733 standard; protein; 412 AA.
AC ABBE64733;
XX
XX 26-MAR-2002 (first entry)
DE Drosophila melanogaster polypeptide SEQ ID NO 20991.
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
OS Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
XX N-PSDB; ABL08836.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signaling and cell-cell
XX interactions.
XX
XX Disclosure; SEQ ID NO 20991; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABLI6176-ABLI30511), expressed DNA
XX sequences (ABLI01840-ABLI6175) and the encoded proteins (ABBS7737-
XX ABB72072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 412 AA;
```

```
Query Match 14.9%; Score 415.5; DB 4; Length 412;
Best Local Similarity 28.7%; Pred. No. 3.7e-25;
Matches 117; Conservative 80; Mismatches 174; Indels 37; Gaps 14;

Qy 14 AHSKMKVYEWKHLDPDGSDEHRAAIKSGEPDHTKNYPFDVD--RWRDKT---FVTIE 67
Db 19 AQALSVFGAYNLELEFFPQERQALRDGLYDPGSVIPIDVDVYKKGDATEPSIFVTIP 78
Qy 68 R-NNGVPSLLVNVTKKGGKGPLLRYP--DWSFAKYEDCSGIVSAFKIADVDFDLWL 124
Db 79 RPAKGVPSYLVVYVNMERNPNTLLQAYPSYEWKSHGADCNGLTSVYRTQIDECGRWIL 138
Qy 125 DSGLVNNQPMGCSKLLTDFDLTKTSKLVKQVEIPHNTAVNATTCMGELVSLVAQAI DRTNT 184
Db 139 DSGEIDFIQ-HCPPQLYAI DLESGKVAHQVKKPKRL---YKGVSRFVPTTVE-LDPHNC 193
Qy 185 ---MVYIADEKGBGLIMYQNSDDSFHRLTSNTFDY-DPRYTKLTAVAGESFTVKNGICGIA 240
Db 194 DVGFTVMADSIGDGIYVVDVAAQQSWRI-ENKFTYPHPDFGTTIAGESFQLMDGVSTT 252
Qy 241 LSP----VTNNLYYSPLSHGLYVYDTEQFRN-PQYEENNVQYEGSQDIL-----NTQSGF 291
Db 253 LTPHGLGGRMMYFHSLSSEWQMAIPLDVVNGSNWRLNDVSAALDQFQLLGRGSCQVA 312
Qy 292 KVVSKNGVLFGLVNGSGIACVNEHVLQRESFDVVAQBEETLQMVSMKIMENLPQSGR 351
Db 313 AAMSESGFLTCGLVQVAPASLLAWNIRTGYSHONLVMLEVEDEQRIQFASGLKIVRN----- 366
Qy 352 INDPEGNEYMLALSNEWMOKIINNDFNFVNFRIILGANVDDLMRNTRC 399
Db 367 ---HEGKEBELVLSNRLQKAFGAGLDYKEINFRIQKCGVQELLUSGRPC 411

RESULT 13
ABBE64727
ID ABBE64727 standard; protein; 415 AA.
XX
XX ABBE64727;
XX
XX 26-MAR-2002 (first entry)
DE Drosophila melanogaster polypeptide SEQ ID NO 20973.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
XX N-PSDB; ABL08830.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signaling and cell-cell
XX interactions.
XX
XX Disclosure; SEQ ID NO 20973; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
```







GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2006, 19:01:48 ; Search time 39.3539 Seconds  
(without alignments)  
1281.133 Million cell updates/sec

Title: US-10-525-567-3

Perfect score: 2796

Sequence: 1 AAVNHQKSAANLAHSMKVI.....NNQNDNRNDNQVHSSKLH 524

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_80.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2269	81.2	467	2 S39193	royal jelly protei
2	1347	48.2	464	2 S39194	royal jelly protei
3	427	15.3	541	2 A25696	yellow protein - f
4	418.5	15.0	453	2 JC7253	Yellow-B protein -
5	278.5	10.0	3848	2 T17414	TipC protein - ali
6	267	9.5	2150	2 S71629	sensory transducti
7	261	9.3	1712	2 C71618	hypothetical prote
8	251	9.0	2819	2 T09080	probable chloroqui
9	247	8.8	568	2 JC7210	molluscan shell ma
10	247	8.8	1584	2 T18276	protein-tyrosine k
11	241	8.6	448	2 S05355	hypothetical prote
12	239.5	8.6	518	2 S23692	erythrocyte membra
13	236.5	8.5	1390	2 T14004	trfA protein - ali
14	227.5	8.1	4550	2 T18440	hypothetical prote
15	226.5	8.1	451	2 A23535	clustered asparagi
16	223.5	8.0	2452	1 RNZQ2L	DNA-directed RNA p
17	222.5	8.0	1650	2 T18444	hypothetical prote
18	222	7.9	1711	2 T18429	hypothetical prote
19	221.5	7.9	719	2 S61046	ARF1 protein - yea
20	220	7.9	954	1 S20907	endo-1,4-beta-xyla
21	219.5	7.9	1093	2 T18275	1-phosphatidylinos
22	218.5	7.8	234	2 S14469	asparagine-rich pr
23	218.5	7.8	1817	2 D71606	hypothetical prote
24	217.5	7.8	537	2 A23770	asparagine-rich pr
25	217.5	7.8	3394	2 T18501	hypothetical prote
26	217	7.8	720	2 T51007	hypothetical prote
27	215.5	7.7	2457	2 T18492	hypothetical prote
28	215	7.7	3844	2 T18402	asparagine/asparta
29	212.5	7.6	725	2 T03219	G-quartet DNA bind

30	210	7.5	2500	2 G71609	hypothetical prote
31	209	7.5	839	2 H90577	lipoprotein vsai l
32	209	7.5	1256	2 S14556	asparagine-rich pr
33	208.5	7.5	1619	2 T18499	hypothetical prote
34	207	7.4	431	2 S50977	hypothetical prote
35	206.5	7.4	947	2 T08605	hypothetical prote
36	206.5	7.4	964	2 S48404	probable membrane
37	205.5	7.3	1254	2 T18277	kinesin heavy chai
38	201.5	7.2	1188	2 A71621	protein with 5'-3'
39	201	7.2	490	2 A46391	CAMP receptor subt
40	199.5	7.1	1278	2 A71609	probable secreted
41	199	7.1	608	2 T18437	hypothetical prote
42	196.5	7.0	1844	2 D71612	hypothetical prote
43	196	7.0	686	2 A71607	Mtn3/RAGLIP-like p
44	193.5	6.9	239	2 S14470	asparagine-rich pr
45	193	6.9	2391	2 T18410	carbamoyl-phosphat

ALIGNMENTS

RESULT 1

S39193

royal jelly protein RJP57-1 - honeybee

C;Species: Apis mellifera (honeybee)

C;Date: 05-Mar-1995 #sequence\_revision 12-May-1995 #text\_change 12-May-1995

C;Accession: S39193

R;Klaudiny, J.; Hanes, J.; Kulifajova, J.; Albert, S.; Simuth, J.

A;Submitted to the EMBL Data Library, September 1993

A;Description: Molecular cloning of two cDNAs from the head of the nurse honey bee (Apis

A;Reference number: S39193

A;Accession: S39193

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-467 <KLA>

A;Cross-references: UNIPARC:UPI0000179A48; EMBL:Z26318

Query Match 81.2%; Score 2269; DB 2; Length 467;  
Best Local Similarity 96.9%; Pred. No. 2.7e-142;  
Matches 435; Conservative 1; Mismatches 11; Indels 2; Gaps 2;

Qy	1	AAVNHQKSAANLAHSMKVIYEWKHIDFDGSDRRDAAIKSGBEDHTKNTYPPFDVDRWD	60
Db	21	AAVNHQKSAANLAHSMKVIYEWKHIDFDGSDREMLRLNL-RIDHTKNT-FDVDRWD	78
Qy	61	KTFTVIERNGVPSLLNVVTKKGGPILLRPYPDMGFAKYEDCGSVSAFKIADVDFDR	120
Db	79	KTFTVIERNGVPSLLNVVTKKGGPILLRPYPDMGFAKYEDCGSVSAFKIADVDFDR	138
Qy	121	LWVLDSGLVNNQPMCSFKLLTFDLKTSKLVQKVEIPHNIAVNATTGMGELVSLAVQAID	180
Db	139	LWVLDSGLVNNQPMCSFKLLTFDLKTSKLVQKVEIPHNIAVNATTGMGELVSLAVQAID	198
Qy	181	RTNTWVYIADEKGGLIMYQNSDDSFHRLTSNTFDYDPRYTKLVAGSSFTVQKIGCIA	240
Db	199	RTNTWVYIADEKGGLIMYQNSDDSFHRLTSNTFDYDPRYTKLVAGSSFTVQKIGCIA	258
Qy	241	LSPTVNNLYSPSSHGLIYYVDTQFRNPQYEEENNVOYEGSQDILNTOSFGKVVSKGVL	300
Db	259	LSPTVNNLYSPSSHGLIYYVDTQFSNPQYEEENNVOYEGSQDILNTOSFGKVVSKGVL	318
Qy	301	FLGLVGNSGIACVNEHQLQRESFDVVAQNETLQMIVSMKIMENLPQSGRINDEGNEY	360
Db	319	FLGLVGNSGIACVNEHQLQRESFDVVAQNETLQMIVSMKIMENLPQSGRINDEGNEY	378
Qy	361	MLALSNRMQKIINDDFNFDVNFRLGANVDDLMENTRCGRYHQNAGNQADNQADNQ	420
Db	379	MLALSNRMQKIINDDFNFDVNFRLGANVDDLMENTRCGRYHQNAGNQADNQADNQ	438
Qy	421	NANNQADNQANQNGNRQNDNRQNDNK	449
Db	439	NANNQADNQANQNGNRQNDNRQNDNK	467



Db 81 SLAYLDLNDTSSKGPALKFPFQAHNLQEAPELVSPFRVADRCGLRWLSDRISGVL 140  
Qy 130 NNNQPMCSPLKLLFDLTKSLVKQVEIPHNIANNATTGMEGLVSLAVQALDRTNTWYIA 199  
Db 141 EOTKIYGAQLLVYDLHNDLLRRHVLPAQLKGSL-----LANLAVEDSDCENTFAYAA 196  
Qy 190 DEKGEGLIMYQNSDDSFHRLTNTDFDPRYTKLTVAGSFTVKGICGIALS-PVTN-- 246  
Db 197 DLGSPGLVYVSWKDEESWRVQHFFHDPDMAGNPSINGIEFQDDGGLYGLALSKPLETGY 256  
Qy 247 -NLVYSLSGHLYVYVTEQFRNPQYBENNQYB-----GSQDILNTQSFQKVSKN-GV 299  
Db 257 ATLFFHPLCSTTESVDTSLRNKTLATSPMIYREFKVLGSRG-PNTQAGAEFLDPDTGV 315  
Qy 300 LFLGLVNGSIACVNEHOVQRESFDVVAQNEETLQIMVSMKIMENL PQSGRINDPEGNE 359  
Db 316 LFVALPNLNEVACWRTATDFSHSQSRHNMNDTLVFPFSIKV-----DDQK 362  
Qy 360 YMLALSRMOKIINDDFNDFNVFRILGANVDDLNRTRCGRYHQNAGNQADNQADN 419  
Db 363 RLWVLSNQLPVFIYDELYAGSINFRILTASVKEAIENTAC---EIRTSPLPDVINKLGD 419  
Qy 420 QNANNQADNQANQON 436  
Db 420 LNTNKLKSNASSLRN 436  
RESULT 5  
TipC protein - slime mold (Dictyostelium discoideum)  
C:Species: Dictyostelium discoideum  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T17414  
R:Stege, J.T.; Laub, M.T.; Loomis, W.F.  
A:Submitted to the EMBL Data Library, July 1998  
A:Description: Interaction of tip genes in early Dictyostelium discoideum development.  
A:Reference number: Z18774  
A:Accession: T17414  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-3848 <STE>  
A:Cross-references: UNIPROT:O76737; UNIPARC:UPI0000076F0B; EMBL:AF079445; NID:G3420744;  
A:Experimental source: strain AX4  
C:Genetics:  
A:Gene: tipC  
A:Introns: 72/3  
C:Function:  
A:Description: required for tip formation  
Query Match 10.0%; Score 278.5; DB 2; Length 3848;  
Best Local Similarity 20.7%; Pred. No. 1.3e-09;  
Matches 118; Conservative 84; Mismatches 226; Indels 143; Gaps 19;  
Qy 68 RNNGVPSSLVNVTNKKGGPGL-----RYP-----DWSFAKYEDCS--- 105  
Db 1737 RNILPLSLKSNKREKQKQEQILPQIIIGSIKPPPIQOSLPQISIDGMFIFVPHSHIP 1796  
Qy 106 ---GVSAFKIADKFDRLWLDSGLVNNQPMCSFKLLTFDLTKSLVKQVEIPHNTAV 162  
Db 1797 VQDSIMKLSKFEATEAWNYTIKVYGEAPNDEIIQSGLSLSIKDI--KISIPGDCV 1854  
Qy 163 NA-----TTGMEGLVSLAVQALDRTNTWY-----IADEKGEGLIMYQN 201  
Db 1855 NSGQGENEYPLFIRSLDIQTTSRGAMGVETITLDANNIQRNQLVDVTNNENNNNN 1914  
Qy 202 SDDSFRHTSNTDFDPRYTKL-----TVAGESFTVK 233  
Db 1915 NNLTTSTTTTTNNQSNKIIAPFKIVQTTSTLLNSESLSLNNIVTFTSYQDFKLM 1974  
Qy 234 NGICGIALSPVTNNLYSPSSHGL--YYVD-----TRQFRNPQYEEENVQYE 279  
Db 1975 MKIINSIIESNNQNLIQEKKERNYLLKRYDPNECTSDDELEBQLQLENDYNDNNYSN 2034

Qy 280 GSQDILNTQSFQKVS-KVNGVL--FLGLVNGSIACVN-EHQV-----LQRE 322  
Db 2035 GNNN--NNNSNNQLPVKESQQLKYLISLEKGEFLINDHDKISSPIKLLSIGVGLKSN 2092  
Qy 323 SPDVVAQNEETLQIMVSMK-----IMENLPQSGRINDPEGNEYMALSNRMQ 369  
Db 2093 IFSFPQKQIQTALSLDANMKAGYFNKIGIWEPIENWGFSTNNSIEGGMVNFNSKIP 2152  
Qy 370 KIINDDFNDFNVFRILGANVDDLNRTRCGRYHQNAGNQADNQADNQANN----- 424  
Db 2153 LYINITKIFIDTSISTYQIWADE-----YYSQ-----QKDKKKNKSYNDNDEIED 2199  
Qy 425 --QVADNQANQKNGRQNDNRQNDKQNGRQNDKQNGRQNDKQNGRQNDKQNGRQNDKQND 482  
Db 2200 TVSQVANDLIKPADDOCKNNNNKNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2259  
Qy 483 NKQNGRQNDKQNGRQNDNRQNDKQNGRQNDKQNGRQNDKQNGRQNDKQNGRQNDKQND 513  
Db 2260 NNNNNNNKNN 2289  
RESULT 6  
S71629  
sensory transduction histidine kinase dhkA - slime mold (Dictyostelium discoideum)  
C:Species: Dictyostelium discoideum  
C:Date: 29-Jan-1998 #sequence\_revision 06-Feb-1998 #text\_change 09-Jul-2004  
C:Accession: S71629  
R:Wang, N.; Shaulsky, G.; Escalante, R.; Loomis, W.F.  
EMBO J. 15, 3890-3898, 1996  
A:Title: A two-component histidine kinase gene that functions in Dictyostelium developme  
A:Reference number: S71629; MUID:96324397; PMID:8670894  
A:Accession: S71629  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-2150 <WAN>  
A:Cross-references: UNIPROT:Q23863; UNIPARC:UPI000017B18D; EMBL:U42597  
A:Experimental source: strain AX4  
C:Genetics:  
A:Gene: dhkA  
A:Map position: 6  
A:Keywords: adenylation; phosphorylation; phosphoprotein; phosphotransferase; two-component regul  
F:2076/Binding site: phosphate (Asp) (covalent) #status predicted  
Query Match 9.5%; Score 267; DB 2; Length 2150;  
Best Local Similarity 20.2%; Pred. No. 3.3e-09;  
Matches 126; Conservative 64; Mismatches 207; Indels 226; Gaps 21;  
Qy 4 NHQRKSANNLAHSMKVIYEWKHIDFEGSDERR-----DAAIKGGEFDPHTKXYPFD- 54  
Db 235 NNGNNNNNITDPTKSKRHSYETNIGSHQRRKSIQSLIANSIHS--PSKLKNKPLSS 292  
Qy 55 -----VDRWRDKTFVTIER-----NNGVPSL 76  
Db 293 STPSTVATCGAVNN 351  
Qy 77 NVVTNKKGKGLLRYPYDMSFAKYEDCSGVSAFKIADKFDRLWLDSGLVNN--NQ 134  
Db 352 NCGNCGSNGGIPPLSP-----RNLSLNSGVNVSPRIHL-----NNLNNSNLP 397  
Qy 135 MCSPKLLTFDLTKSLVKQVEIPHNIANNATTGMEGLVSLAVQALDRTNTWYIADEKGE 194  
Db 398 PLSPRHINFINVSNLNN-----NNNNINPNNP----- 427  
Qy 195 GLIMYQNSDDSFHRLT--SNTFDYDPRYTKLTVAGESFTVKGICGIALSPVTNNLYSP 252  
Db 428 ----NNSNNNNNNVSPRNNNNISPR-----GSNISPRNNGSGSTT 464  
Qy 253 LSSHGLYVYDTEQFRNPQYBENNQYBENNQYBENNQYBENNQYBENNQYBENNQYBENN 308  
Db 465 ISP-----RNTSNNNNIINNINNNILTPPRNSPRLENVNPPTNSPRLLATSLNS 513  
Qy 309 GIACVNEHQVLRQSFVVAQNEETLQIMVSMKIMENL PQSGRINDPEGNEYMALSNRM 368







C;Keywords: surface antigen

```
Query Match      8.6%; Score 239.5; DB 2; Length 518;
Best Local Similarity 20.3%; Pred. No. 2.8e-08;
Matches 99; Conservative 74; Mismatches 137; Indels 177; Gaps 20;

Qy 128 LVNNQPMCSFKLLTFDLKTSKLVKQVEIPHNIATVATTGMEGLVSLAVQAID--RNTM 185
Db 75 MMSNENKMDPKHVENLQNNILKR-----DNIIRKPRVKLSNLLFGVDIKLVDM 130
Qy 186 VYIADEKGEGLIMYONSDGFFHLRTSNTFDYPRYTKLTVAGSFTVKNGICGIALSP-- 243
Db 131 TFYNDWN---LFPKDTGMP-----GLAGKELSCRN-----ISAWPNI 164
Qy 244 -VTNNLYYS-----PLSSHGLYVVDTEQFRNPQYEE-----NNV----- 276
Db 165 LENNSIFALDINLPNDIAIKTYPTFPVNSVGTLYEYDLKKGDNKKSQDQINNVKDNM 224
Qy 277 -----QYEGSQDILN-----TQSPGKVSNGVFLGLVGNSGIA-CVNEHQ 317
Db 225 IDNKNVDVGQVQINIEFETSSVSIHIGNSKEYGNIIMSEPRERGIFDSFIQFTIYNLE 284
Qy 318 VLQRESFDVVAQNEETL-----QMIVSMK-----IMENLPQSGRINDPEGNEYMALSNR 367
Db 285 VCNKNIIFGKYSNWQAIIDLSQQCLVLPKFWLSIMEYLP----- 324
Qy 368 MQKIINNDPNFNVNFRILGANVDLDMRTRCGRYHQNAGNQADNQADNQANNQNA 427
Db 325 ----VNKD-----DD-----RCIR-KMNFNNNNNNKNDNNNNNNKNN 359
Qy 428 DNQVANKQNGRQNDNQKNGRQNDNQKNGRQNDNQKNGRQNDNQKNGRQNGKQ----- 480
Db 360 DNNN-----NNKNNDNNNNNNKNDNNNNNNKNDNNNNNNKNDNNNGDNDVDYEENSLP 415
Qy 481 -----NDNKQNGRQNDKNGRQNDNQ-----NNQND 509
Db 416 RMCSVDHNRPLPLTKFPLSDNDIVSDNNIYSSPESKNGKNEKQAIYIPLDMLII 475
Qy 510 NNRNDNQ 516
Db 476 NDKKNQ 482

RESULT 13
T14004
trfA protein - slime mold (Dictyostelium discoideum)
C;Species: Dictyostelium discoideum
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T14004
R;Salto, J.; Adachi, H.; Sutoh, K.
J. Biol. Chem. 273, 24654-24659, 1998
A;Title: Dictyostelium TRFA homologous to yeast Ssn6 is required for normal growth and
A;Reference number: Z17852; MUID:98406112; PMID:9733762
A;Accession: T14004
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1390 <SAI>
A;Cross-references: UNIPROT:O77033; UNIPARC:UPI000007B414; EMBL:AB009080; NID:d1228566;
A;Gene: trfA
A;Introns: 333/3; 364/3; 637/1

Query Match      8.5%; Score 236.5; DB 2; Length 1390;
Best Local Similarity 21.7%; Pred. No. 1.8e-07;
Matches 94; Conservative 42; Mismatches 128; Indels 169; Gaps 17;

Qy 239 IALSPVTNNLYYSPLSHGLYVVDTEQF-----RNPQVEENNVOYEGSQDILNTQSF 290
Db 481 IRLNPLFLSEVMY-----DLGTLYESCHQHTDSLDAYQAAELDPNHKIQSLRLATLRAQVS 536
Qy 291 GKVSVKNG----- 309
Db 537 GKPIGKDGYDLONGEHEGKGKSTPMIIEPNSPQTGAMESLGKGGQNNRNGNNNNNSFV 596
```

```
Qy 310 --IACVNEH-QVLQRESFDVVAQNEETLQMI-----VSMKIM 343
Db 597 PELAEINSHLPEDMRNNSLSPSSNELNSLIDRKTGERGEDMHNHSHHQSYSMSMNMM 656
Qy 344 ENLPQSGRINDPEGNEYMALSNRMOKIINNDFNFNDV--NFRILGANVDDDLMTNRCGRY 402
Db 657 NNNNNNNNN-----NMNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 708
Qy 403 --HNQAGNQADNQADNQANNQADNQANKQNG-----KONGRQND 467
Db 709 NNHQMNQYNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 767
Qy 438 -NRQN--DNQON-----DNKO-----NGNRQNDN-----KONGRQND 467
Db 768 DNRDNRDNRDNRDNRDNRDNRDNRDNRDNRDNRDNRDNRDNRDNRDNRDNRDNRDNR 827
Qy 468 NKONGNRQNGKNDKNDKQNG-----NRQNDKNGRQND 502
Db 828 NYNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 887
Qy 503 NQNNQNDNRRNDN 515
Db 888 NNNNNNNNNNN 900

RESULT 14
T18440
hypoetical protein C0425w - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T18440
R;Lawson, D.; Bowman, S.; Barrell, B.
submitted to the EMBL Data Library, August 1997
A;Reference number: Z18935
A;Accession: T18440
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-4550 <LAW>
A;Cross-references: UNIPROT:O77336; UNIPARC:UPI0000110116; EMBL:Z98547; NID:e1325376; PFI
C;Genetics:
A;Map position: 3
A;Note: C0425w

Query Match      8.1%; Score 227.5; DB 2; Length 4550;
Best Local Similarity 18.5%; Pred. No. 3.8e-06;
Matches 123; Conservative 77; Mismatches 204; Indels 262; Gaps 22;

Qy 69 NNGVPSSLNVVT---NKGKGGPLLRP-----YPDWSFAKYEDCSGIVSAFKIA 114
Db 1529 NNNNISNNNVQEDYKMKGKSGHKQKHKSTNYNQDDIYVNDPSSI--CSLLSSSLSTN 1586
Qy 115 VDKFDRLLWLDLGLVNNNQPMCSFKLLT---FDLKTSLKLVKQVEIPHNIATVATTGME 170
Db 1587 ND-----IDSSYVSNNSLYNSSSMFSDNDPCLCYSSVSSSYEYDDINNVLNKNKIHN 1639
Qy 171 LVSLAVQAIDRTVMYVIADKGEGLIMYONSDGFFHLRTSNTFDYDPRY-----TKLTV 225
Db 1640 FPSFYKNELNNNN-----NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1684
Qy 226 AGESF--TVKNGICGIALSPVTNNL-----YYSPLSHGLYVVDTEQFRNPQYEE 274
Db 1685 SRSRFGWSLYNTSDNISSEISNNLSFSKSNYCSSISDHEI--EKDVTIYRNTVHRII 1743
Qy 275 NVQYEG-----SQDILNTQSPGKVSNGVFLGLVGNSGIACVNEHQVLR----- 321
Db 1744 NRWKDKKNTGSKKRDMSNKKGDKSCNKKN-----ICDNKKNIYDNKKLIYD 1790
Qy 322 -----ESPDVVAQNEETLQMIYSMKIME-NLPQSGRINDPEGNEYMAL----- 364
Db 1791 NKDSSTDAPVENIDHLIEKEEIDNILLNLGLTYNLK-----NNNNNNNNNNNNNNNNNN 1845
Qy 365 --SNR--MQKIIN-----NDFNFNDV-----FRILGANVD 391
```



```

Db      1846  THSNEHDTKAVMIKILNTLNKKYNNINCIFYHYNSFLHSFLSYDEYIELIKERKED 1905
Qy      392  DLARNTRCGRYHQNAGNQADNQANQONQONQONQONQONQONQONQONQONQON 451
Db      1906  IEKEIKRGKYCHQEDRDDNNNNNDNNNNNDNNNNNDNNNNNDNNNNNDNNNNNDNNNS 1965
Qy      452  G----- 452
Db      1966  GEDQLINEDNKRKINNLSHETHNNMSDHNKPKKKKKNSNETQINDDTNKKETHKNEI 2025
Qy      453  -----NRQND----- 457
Db      2026  DHKEQKQDTCNGIEEKKDQVSKSNLIISKVFHFFERKNNNIKKDEHKKYNERNDKQSE 2085
Qy      458  -----NKQNGNRQNDKQNGNRQNGKQNG--DNKQNGNRQNDKQNGNRQNDKQNGNRQND 509
Db      2086  CDMLLQNKRDNDHNNNNNNDDNNNNNNDDNNNNNNDDNNNNNNDDNNNNNNDDNNNNNN 2145
Qy      510  NNRNDN 515
Db      2146  NNDDDN 2151

RESULT 15
A23535
clustered asparagine-rich merozoite-associated antigen - malaria parasite (Plasmodium fa
C;Species: Plasmodium falciparum
C;Date: 02-Jun-1988 #sequence_revision 02-Jun-1988 #text_change 09-Jul-2004
C;Accession: A23535; A49245
R;Wahlgren, M.; Aslund, L.; Franzen, L.; Sundvall, M.; Wahlin, B.; Berzins, K.; McNicol,
Proc. Natl. Acad. Sci. U.S.A. 83, 2677-2681, 1986
A;Title: A Plasmodium falciparum antigen containing clusters of asparagine residues.
A;Reference number: A23535; MUID:86206015; PMID:3517875
A;Accession: A23535
A;Molecule type: DNA
A;Residues: 1-451 <WAH>
A;Cross-references: UNIPROT:P13824; UNIPARC:UPI00000126056; GB:M13021; NID:g160079; PIDN:
R;Sjolander, A.; Stahl, S.; Lovgren, K.; Hansson, M.; Cavelier, L.; Wallen, A.; Helmdy,
Exp. Parasitol. 76, 134-145, 1993
A;Title: Plasmodium falciparum: the immune response in rabbits to the clustered asparagi
A;Reference number: A49245; MUID:93202225; PMID:8454022
A;Contents: 7G8
A;Accession: A49245
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 9-25 <SJO>
A;Cross-references: UNIPARC:UPI0000078536; GB:S56699; NID:g298738; PID:g298739
A;Note: sequence extracted from NCBI backbone (NCBIN:127859, NCBI:P:127860)

Query Match      8.1%; Score 226.5; DB 2; Length 451;
Best Local Similarity 25.5%; Pred. No. 1.7e-07;
Matches 77; Conservative 46; Mismatches 110; Indels 69; Gaps 13;

Qy      278  YEGSQDILNTQSGFKVSK-----NGVFLGLVGNNGI---AC-----VNEH 316
Db      7  YKMNNEIVNSTIDGLVTLKLIQNIPIPHITDVHLSLLGNVGFICIDYFNKSKKQMMNN 66
Qy      317  QVLQRESFDVVA---QNEETLQIVSMKIMENLPQSGRINDPEGNEYMLALSINRMQKI- 371
Db      67  NFANKKIYNTALVTNTTHEALNVLKNYK-----NLIDTSGEERNIDAKFAPNVYS 117
Qy      372  INNDFNFNDVNFRIIGANVDDILMRNTRCGR-----YHNQN-AGNQADNQADNQNA 422
Db      118  INNNNNNNSNTFFQKN---MNTNFSQGSTNYGNSYNSFQGNMNNMNNYFNYS 173
Qy      423  NNQONADNQANQONG--NRQNDNRQNDKQNGNRQDKQ-----NGNRQNDKQNGNRQ 475
Db      174  SNNNNNQ-TNTQNNFMFRNKKNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 232
Qy      476  NGNKQNDKQNGNRQNDN-----KRNQNGNRQNDNRQNDNRQNDNRQNDNRQNDNRQND 522
Db      233  NFMNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 292

```

Qy 523 LH 524  
Db 293 SH 294

Search completed: May 3, 2006, 19:10:15  
Job time : 42.3539 secs

This Page Blank (Page 3)

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2006, 18:54:38 ; Search time 242.593 Seconds  
(without alignments)

1523.941 Million cell updates/sec

Title: US-10-525-567-3

Perfect score: 2796

Sequence: 1 AAVNHQRKSNLAHSMKVI.....NNQNDNRNDNQVHSSKLLH 524

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 05.80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2773	99.2	544	1 MRJP3 APIME	Q17060 apis mellif
2	1671.5	59.8	468	2 Q5VK55 APICC	Q5vk55 apis cerana
3	1658.5	59.3	463	2 Q6MMJ0 APICE	Q6mmj0 apis cerana
4	1658.5	59.3	463	2 Q5OH66 APICE	Q5oh66 apis cerana
5	1634.5	58.5	452	1 MRJP2 APIME	Q77061 apis mellif
6	1572.5	56.2	437	2 Q6W3E3 APIME	Q6w3e3 apis mellif
7	1486.5	53.2	443	2 Q6IMJ9 APIME	Q6imj9 apis mellif
8	1437.5	51.4	579	2 Q6QN15 APICE	Q6qn15 apis cerana
9	1427	51.0	598	2 Q5VK56 APICC	Q5vk56 apis cerana
10	1420.5	50.8	432	1 MRJP1 APIME	Q18330 apis mellif
11	1420.5	50.8	432	2 Q548D6 APIME	Q548d6 apis mellif
12	1407	50.3	598	1 MRJP5 APIME	Q97432 apis mellif
13	1403.5	50.2	433	2 Q5VLZ2 APICC	Q5vle2 apis cerana
14	1402.5	50.2	433	2 Q6MMJ1 APICE	Q6mmj1 apis cerana
15	1402	50.1	485	2 Q6QN16 APICE	Q6qn16 apis cerana
16	1347	48.2	464	1 MRJP4 APIME	Q17061 apis mellif
17	1218	43.6	416	2 Q6TGR0 APIME	Q6tgr0 apis mellif
18	1132.5	40.5	423	2 Q4ZJX1 APIME	Q4zjx1 apis mellif
19	834	29.8	222	2 Q5I222 APICE	Q5i222 apis cerana
20	819	29.3	220	2 Q5I223 APICE	Q5i223 apis cerana
21	778.5	27.8	222	2 Q5I224 APICE	Q5i224 apis cerana
22	490.5	17.5	186	2 Q6EM24 APIDO	Q6em24 apis dorsat
23	463.5	16.6	406	2 Q7Q4T5 ANOAGA	Q7q4t5 anopheles g
24	454.5	16.3	463	2 Q9V4C0 DROME	Q9v4c0 drosophila
25	451	16.1	411	2 Q6DLY9 APIME	Q6dly9 apis mellif
26	444	15.9	461	2 Q7PHB5 ANOAGA	Q7phb5 anopheles g
27	433.5	15.5	432	2 Q9W1R1 DROME	Q9w1r1 drosophila
28	433	15.5	568	1 YELL_DROU	O02437 drosophila
29	431	15.4	568	1 YELL_DROMD	Q9gp71 drosophila
30	429	15.3	541	1 YELL_DROYA	Q9bi17 drosophila
31	428	15.3	568	1 YELL_DROGU	Q9gp81 drosophila

32	427	15.3	541	1 YELL_DROMA	P62408 drosophila
33	427	15.3	541	1 YELL_DROME	P09957 drosophila
34	427	15.3	541	1 YELL_DROSI	P62407 drosophila
35	426	15.2	462	1 Q7PMG1 ANOAGA	Q7pmg1 anopheles g
36	426	15.2	497	2 Q7Z0H1_9DIPT	Q7z0h1 drosophila
37	426	15.2	540	2 Q86CU0_9DIPT	Q86cu0 drosophila
38	425	15.2	541	2 Q8ISG5_DROVI	Q8isg5 drosophila
39	423	15.1	494	2 Q7Z0H2_9DIPT	Q7z0h2 drosophila
40	422.5	15.1	426	2 Q9VVF2 DROME	Q9vfv2 drosophila
41	422.5	15.1	546	2 Q5IBN4_9DIPT	Q5ibn4 drosophila
42	419	15.0	541	1 YELL_DROER	Q9bi23 drosophila
43	418.5	15.0	453	2 Q9VJIS_DROME	Q9vjis drosophila
44	417	14.9	560	1 YELL_DROPS	Q9bi18 drosophila
45	415.5	14.9	412	2 Q9W1R0_DROME	Q9w1r0 drosophila

#### ALIGNMENTS

RESULT 1					
MRJP3 APIME					
ID	MRJP3 APIME	STANDARD;	PRT;	544 AA.	
AC	Q17060;				
DT	28-FEB-2003 (Rel. 41, Created)				
DT	28-FEB-2003 (Rel. 41, Last sequence update)				
DT	10-MAY-2005 (Rel. 47, Last annotation update)				
DE	Major royal jelly protein 3 precursor (MRJP-3) (Bee-milk protein)				
DE	(Royal jelly protein RJP57-1).				
GN	Name=MRJP3;				
OS	Apis mellifera (Honeybee).				
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
OC	Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;				
OC	Apidae; Apis.				
OX	NCBI_TaxID=7460;				
RN	[1]				
RP	NUCLEOTIDE SEQUENCE.				
RC	TISSUE=Head;				
RA	Klaudiny J., Hanes J., Kulifajova J., Albert S., Simuth J.;				
RT	"Molecular cloning of two cDNAs from the head of the nurse honey bee				
RT	(Apis mellifera L.) for coding related proteins of royal jelly.";				
RL	J. Apicultural Res. 33:105-111(1994).				
RN	[2]				
RP	SEQUENCE REVISION TO THE C-TERMINUS.				
RA	Albert S., Klaudiny J., Simuth J.;				
RT	"Newly discovered features of the updated sequence of royal jelly				
RT	protein RJP571; longer repetitive region on C-terminus and homology to				
RT	Drosophila melanogaster yellow protein.";				
RL	J. Apicultural Res. 35:63-68(1996).				
RN	[3]				
RP	NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 21-36.				
RC	TISSUE=Head;				
RX	MEDLINE=99007754; PubMed=9791542; DOI=10.1007/s000180050229;				
RA	Schmitzova J., Klaudiny J., Albert S., Schroeder W., Schreckengost W.,				
RA	Hanes J., Judova J., Simuth J.;				
RT	"A family of major royal jelly proteins of the honeybee Apis mellifera				
RT	L.";				
RL	Cell. Mol. Life Sci. 54:1020-1030(1998).				
RN	[4]				
RP	NUCLEOTIDE SEQUENCE, AND PARTIAL PROTEIN SEQUENCE OF 45-59; 70-77;				
RP	106-120; 214-224 AND 243-253.				
RC	TISSUE=Hypopharyngeal gland;				
RX	MEDLINE=98055707; PubMed=9395329;				
RA	Ohashi K., Natori S., Kubo T.;				
RT	"Change in the mode of gene expression of the hypopharyngeal gland				
RT	cells with an age-dependent role change of the worker honeybee Apis				
RT	mellifera L.";				
RL	Eur. J. Biochem. 249:797-802(1997).				
CC	-!- FUNCTION: May play an important role in honeybee nutrition. It is				
CC	found in the royal jelly which is the food of the queen honey bee				
CC	larva. The royal jelly determines the development of the young				
CC	larvae and is responsible for the high reproductive ability of the				
CC	honeybee queen.				
CC	-!- SUBCELLULAR LOCATION: Secreted.				

CC -1- TISSUE SPECIFICITY: Hypopharyngeal glands of nurse honey bees.  
 CC -1- DEVELOPMENTAL STAGE: Produced by the cephalic glandular system of  
 CC the nurse honey bee.  
 CC -1- SIMILARITY: Belongs to the major royal jelly protein family.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 CC EMBL; Z26318; CAA81227.1; -; mRNA.  
 CC DR Ensembl; ENSAPMG0000007331; Apis mellifera.  
 CC DR InterPro; IPR003534; Royaljelly.  
 CC DR PANTHER; PTHR10009; Royaljelly; 1.  
 CC DR Pfam; PF03022; MRJP; 1.  
 CC DR PRINTS; PR01366; ROYALJELLY.  
 CC KW Direct protein sequencing; Glycoprotein; Repeat; Signal.  
 FT SIGNAL 1 20  
 FT CHAIN 21 544 Major royal jelly protein 3.  
 FT REPEAT 424 428 1.  
 FT REPEAT 429 433 2.  
 FT REPEAT 434 438 3.  
 FT REPEAT 439 443 4.  
 FT REPEAT 444 448 5.  
 FT REPEAT 449 453 6.  
 FT REPEAT 454 458 7.  
 FT REPEAT 459 463 8.  
 FT REPEAT 464 468 9.  
 FT REPEAT 469 473 10.  
 FT REPEAT 474 478 11.  
 FT REPEAT 479 483 12.  
 FT REPEAT 484 488 13.  
 FT REPEAT 489 493 14.  
 FT REPEAT 494 498 15.  
 FT REPEAT 499 503 16.  
 FT REPEAT 504 508 17.  
 FT REPEAT 509 513 18.  
 FT REPEAT 514 518 19.  
 FT REPEAT 519 523 20.  
 FT REGION 424 523 20 X 5 AA tandem repeats of [NKR]-Q-N-  
 FT CARBOHYD 183 183 N-linked (GlcNAc...) (Potential).  
 FT SQ SEQUENCE 544 AA; 61662 MW; 4C8FFFC8A2759F52 CRC64;  
 Query Match 99.2%; Score 2773; DB 1; Length 544;  
 Best Local Similarity 99.4%; Pred. No. 4.1e-167;  
 Matches 521; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 Qy 1 AAVNHQKSAANLAHSMKVIYEWKHIDDFGSDERRDAAIKSGEFDHTKNYPFDVDRWRD 60  
 Db 21 AAVNHQKSAANLAHSMKVIYEWKHIDDFGSDERRDAAIKSGEFDHTKNYPFDVDRWRD 80  
 Qy 61 KTFVTIERNNGVPSLNVVTKNGKGGPILRPYPDWSFAKVEDCSGIVSAFKIADVDFDR 120  
 Db 81 KTFVTIERNNGVPSLNVVTKNGKGGPILRPYPDWSFAKVEDCSGIVSAFKIADVDFDR 140  
 Qy 121 LWLDSGLVNNQPMCSFKLLTFDLKTSKLVKQVEIPHNIATVAVNATGMBELVSLAVQVAD 180  
 Db 141 LWLDSGLVNNQPMCSFKLLTFDLKTSKLVKQVEIPHNIATVAVNATGMBELVSLAVQVAD 200  
 Qy 181 RTNTMVIADKEGGLIMYQNSDDSFRLTSNTFDYDPRYTKLTVAGSFVTKNGICGIA 240  
 Db 201 RTNTMVIADKEGGLIMYQNSDDSFRLTSNTFDYDPRYTKLTVAGSFVTKNGIYGIA 260  
 Qy 241 LSPVTNNLYSPSSHGLIYYVDTEQFRNPQYEEENNVOYEGSDILNTQSGKVVSKNGVL 300  
 Db 261 LSPVTNNLYSPSSHGLIYYVDTEQFRNPQYEEENNVOYEGSDILNTQSGKVVSKNGVL 320  
 Qy 301 FLGLVGNSGIACVNEHQVLORESFDVVAQNEETLQIMVSKIMENLPQSGRINDPEGNEY 360  
 Db 321 FLGLVGNSGIACVNEHQVLORESFDVVAQNEETLQIMVSKIMENLPQSGRINDPEGNEY 380

Qy 361 MLALSNRMQKIINNDNFNDFNFRILGANVDDLMRNTRCGRYHQNAGNQADNQ 420  
 Db 381 MLALSNRMQKIINNDNFNDFNFRILGANVDDLMRNTRCGRYHQNAGNQADNQ 440  
 Qy 421 NANNQADNQNANKQNGRQNDNRQNDKQNGRQNDKQNGRQNDKQNGRQNGKQ 480  
 Db 441 NANNQADNQNANKQNGRQNDNRQNDKQNGRQNDKQNGRQNDKQNGRQNGKQ 500  
 Qy 481 NDNKQNGRQNDKQNGRQNDNRQNDKQNGRQNDKQNGRQNDKQNGRQNGKQ 524  
 Db 501 NDNKQNGRQNDKQNGRQNDNRQNDKQNGRQNDKQNGRQNDKQNGRQNGKQ 544  
 RESULT 2  
 QSVK55 APICC  
 ID QSVK55 APICC PRELIMINARY; PRT; 468 AA.  
 AC QSVK55;  
 DT 01-FEB-2005 (TrEMBLrel. 29, Created)  
 DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)  
 DE Major royal jelly protein MRJP2.  
 OS Apis cerana cerana (Oriental honeybee).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;  
 OC Apidae; Apis.  
 OX NCBI\_TaxID=94128;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Su S., Chen S., Albert S.;  
 RL "Molecular cloning of MRJP2 cDNA from Apis cerana cerana in China.";  
 DR EMBL; AY392758; AAR83083.1; -; mRNA.  
 DR InterPro; IPR003534; Royaljelly.  
 DR Pfam; PF03022; MRJP; 1.  
 DR PRINTS; PR01366; ROYALJELLY.  
 SQ SEQUENCE 468 AA; 53060 MW; 5B2A6ARF4C530C84 CRC64;

Query Match 59.8%; Score 1671.5; DB 2; Length 468;  
 Best Local Similarity 66.4%; Pred. No. 1.6e-97;  
 Matches 308; Conservative 74; Mismatches 65; Indels 17; Gaps 3;

Qy 6 QRKSANLAHSMKVIYEWKHIDDFGSDERRDAAIKSGEFDHTKNYPFDVDRWRD 65  
 Db 21 QRKSANLAHSMKVIYEWKHIDDFGSDERRDAAIKSGEFDHTKNYPFDVDRWRD 80  
 Qy 66 IERNNGVPSLNVVTKNGKGGPILRPYPDWSFAKVEDCSGIVSAFKIADVDFDR 125  
 Db 81 ILKYDGVPSLNVVTKNGKGGPILRPYPDWSFAKVEDCSGIVSAFKIADVDFDR 140  
 Qy 126 SGLVNNQPMCSFKLLTFDLKTSKLVKQVEIPHNIATVAVNATGMBELVSLAVQVAD 185  
 Db 141 SGLINRTEPICAPKLHVFDLKNKTKLQIEIPHDIATVAVNATGMBELVSLAVQVAD 200  
 Qy 186 VYIADKEGGLIMYQNSDDSFRLTSNTFDYDPRYTKLTVAGSFVTKNGICGIALSPVT 245  
 Db 201 VYIADKEGGLIMYQNSDDSFRLTSNTFDYDPRYTKLTVAGSFVTKNGICGIALSPVT 260  
 Qy 246 NNLYSPSSHGLIYYVDTEQFRNPQYEEENNVOYEGSDILNTQSGKVVSKNGVL 304  
 Db 261 NNLYSPSSHGLIYYVDTEQFRNPQYEEENNVOYEGSDILNTQSGKVVSKNGVL 320  
 Qy 305 VGNSGIACVNEHQVLORESFDVVAQNEETLQIMVSKIMENLPQSGRINDPEGNEY 364  
 Db 321 VGNSGIACVNEHQVLORESFDVVAQNEETLQIMVSKIMENLPQSGRINDPEGNEY 380  
 Qy 365 SNRMQKIINNDNFNDFNFRILGANVDDLMRNTRCGRYHQNAGNQADNQ 424  
 Db 381 SNRMQKIINNDNFNDFNFRILGANVDDLMRNTRCGRYHQNAGNQADNQ 428  
 Qy 425 QNADNQNANKQNGRQNDNRQNDKQNGRQNDKQNGRQNDKQNGRQNGKQ 468  
 Db 429 QKNNQN-----NNNQNNNQKNNNQKNNNQKNNNQKNNNQNTNN 468

13-SEP-2005 (TReMBLrel. 31, Last annotation update)  
DE Major royal jelly protein 2.  
GN Name=mrjpl;  
OS Apis cerana (Indian honeybee).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;  
OC Apidae; Apis.  
OX NCBI\_TaxID=7461;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Imjongjirak C., Klinbunga S., Sittipraneed S.;  
RT "Cloning, Expression and Genomic Organization of Genes Encoding Major  
RT Royal Jelly Protein 1 and 2 of the Honeybee (Apis cerana).";  
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY515689; AAS88557.1; -; Genomic DNA.  
SQ SEQUENCE 463 AA; 52462 MW; 930C91C8CC595935 CRC64;

Query Match 59.3%; Score 1658.5; DB 2; Length 463;  
Best Local Similarity 69.1%; Pred. No. 1e-96;  
Matches 307; Conservative 62; Mismatches 72; Indels 3; Gaps 2;

Qy 6 QRSANNLAHSMKVIYEWKHI DFGSDERRDAAIKSGEDHTKNYPFDVDRWRDKTFVT 65  
Db 21 RNSAKNLENSLNVHEWKI IDYDFGSEERRQAAIQSGEYDHTKNYPFDVQWHDKTFT 80  
Qy 66 IERNNGVPSSLNVVTKKGGPGLLRYPDPWSFAKYEDCSGIVSAFKIADVDFDLRLWLD 125  
Db 81 ILKYDGVPSLTLMISNKGKGRLLQPPYDWSAENKDCSGIVSAFKIADKDFDLRLWLD 140  
Qy 126 SGLVNNQPMCSKLLTFDLTKSLVKQVEIPHNIAVNATTGMGELVSLAVQAI DRTNTM 185  
Db 141 SGLINRTEPICAPKLHVFDLNTKHLKQIEIPHDI AVNATTGKGLVSLVQVQAMPNTL 200  
Qy 186 VYIADKGEGLIMVQNSDDSFHRLTSNTFDYDPRYKVTINGESFTLKNGICGVALSPVT 245  
Db 201 VYIADHKGDALIVQNSDDSFHRLTSNTFDYDPRYKVTINGESFTLKNGICGVALSPVT 260  
Qy 246 NLYYSPSLSHGLVYVDETFQFRNPQY - EENNVOYEGSDIINTQSFQKVSQNGVLFGL 304  
Db 261 NLYYSPSLASHGLVYVNTPEPFMSQFGDNNNVQYEGSDTLNTQSLAKAVSKDGLVFL 320  
Qy 305 VNSGIACVNEHQVLORESFDVVAQNEETLOMIVSMKIMENLPOSGRINDEPEGNEYMAL 364  
Db 321 VNSALGCLNHPQLQRENLELVAQNEKTLQMIAGMKIKEELPHFVGSNPKVDEYMLVL 380  
Qy 365 SNRMQKIINDFNFDVNVFRILGANVDDLMRNTCRGRYHQNAGNQADNQADNQANN 424  
Db 381 SNRMQKIINDFNFDVNVFRILGANVDDLMRNTCRGRYHQNAGNQADNQADNQANN 438  
Qy 425 QNADNQANKQNGNRQNDNRQNDN 448  
Db 439 QKNNQKNNQKNNQKNNQNTN 462

RESULT 5  
MRJP2 APIME  
ID MRJP2 APIME STANDARD; PRT; 452 AA.  
AC 077061;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-MAY-2005 (Rel. 47, Last annotation update)  
DE Major royal jelly protein 2 precursor (MRJP-2) (Bee-milk protein).  
GN Name=MRJP2;  
OS Apis mellifera (Honeybee).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;  
OC Apidae; Apis.  
OX NCBI\_TaxID=7460;  
RN [1]  
RP NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 18-30.  
RC TISSUE=Head;  
RX MEDLINE=99007754; PubMed=9791542; DOI=10.1007/s000180050229;  
RA Schmitzova J., Klaudiny J., Albert S., Schroeder W., Schreckengost W.,

QBMWJO APICE PRELIMINARY; PRT; 463 AA.  
AC QBMWJO;  
DT 01-OCT-2002 (TReMBLrel. 22, Created)  
DT 01-MAR-2004 (TReMBLrel. 26, Last sequence update)  
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)  
DE Major royal jelly protein MRJP2 precursor.  
GN Name=MRJP2;  
OS Apis cerana (Indian honeybee).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;  
OC Apidae; Apis.  
OX NCBI\_TaxID=7461;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Sittipraneed S., Imjongjirak C.;  
RT "Molecular Cloning of Major Royal Jelly Protein (MRJP2) cDNA from Apis  
RT cerana in Thailand.";  
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF525777; AAM88282.2; -; mRNA.  
DR InterPro: IPR003534; Royaljelly.  
DR Pfam: PF03022; MRJP; 1.  
DR PRINTS; PR01366; ROYALJELLY.  
DR Signal.  
KW SIGNAL.  
FT SIGNAL  
SQ SEQUENCE 463 AA; 52490 MW; 29D92E3749B15992 CRC64;

Query Match 59.3%; Score 1658.5; DB 2; Length 463;  
Best Local Similarity 69.1%; Pred. No. 1e-96;  
Matches 307; Conservative 62; Mismatches 72; Indels 3; Gaps 2;

Qy 6 QRSANNLAHSMKVIYEWKHI DFGSDERRDAAIKSGEDHTKNYPFDVDRWRDKTFVT 65  
Db 21 RNSAKNLENSLNVHEWKI IDYDFGSEERRQAAIQSGEYDHTKNYPFDVQWHDKTFT 80  
Qy 66 IERNNGVPSSLNVVTKKGGPGLLRYPDPWSFAKYEDCSGIVSAFKIADVDFDLRLWLD 125  
Db 81 ILKYDGVPSLTLMISNKGKGRLLQPPYDWSAENKDCSGIVSAFKIADKDFDLRLWLD 140  
Qy 126 SGLVNNQPMCSKLLTFDLTKSLVKQVEIPHNIAVNATTGMGELVSLAVQAI DRTNTM 185  
Db 141 SGLINRTEPICAPKLHVFDLNTKHLKQIEIPHDI AVNATTGKGLVSLVQVQAMPNTL 200  
Qy 186 VYIADKGEGLIMVQNSDDSFHRLTSNTFDYDPRYKVTINGESFTLKNGICGVALSPVT 245  
Db 201 VYIADHKGDALIVQNSDDSFHRLTSNTFDYDPRYKVTINGESFTLKNGICGVALSPVT 260  
Qy 246 NLYYSPSLSHGLVYVDETFQFRNPQY - EENNVOYEGSDIINTQSFQKVSQNGVLFGL 304  
Db 261 NLYYSPSLASHGLVYVNTPEPFMSQFGDNNNVQYEGSDTLNTQSLAKAVSKDGLVFL 320  
Qy 305 VNSGIACVNEHQVLORESFDVVAQNEETLOMIVSMKIMENLPOSGRINDEPEGNEYMAL 364  
Db 321 VNSALGCLNHPQLQRENLELVAQNEKTLQMIAGMKIKEELPHFVGSNPKVDEYMLVL 380  
Qy 365 SNRMQKIINDFNFDVNVFRILGANVDDLMRNTCRGRYHQNAGNQADNQADNQANN 424  
Db 381 SNRMQKIINDFNFDVNVFRILGANVDDLMRNTCRGRYHQNAGNQADNQADNQANN 438  
Qy 425 QNADNQANKQNGNRQNDNRQNDN 448  
Db 439 QKNNQKNNQKNNQKNNQNTN 462

RESULT 4  
Q50H66 APICE PRELIMINARY; PRT; 463 AA.  
AC Q50H66;  
DT 13-SEP-2005 (TReMBLrel. 31, Created)  
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)

```
RA Hanes J., Judova J., Simuth J.;
RT "A family of major royal jelly proteins of the honeybee Apis mellifera
RL L.";
RL Cell. Mol. Life Sci. 54:1020-1030(1998).
CC -|- FUNCTION: May play an important role in honeybee nutrition. It is
CC found in the royal jelly which is the food of the queen honey bee
CC larva. The royal jelly determines the development of the young
CC larvae and is responsible for the high reproductive ability of the
CC honeybee queen.
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- TISSUE SPECIFICITY: Found in the hypopharyngeal glands.
CC -|- DEVELOPMENTAL STAGE: Produced in the cephalic glands of the nurse
CC honeybee.
CC -|- SIMILARITY: Belongs to the major royal jelly protein family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL: AF00632; AAC61894.1; -; mRNA.
CC Ensembl: ENSAPMG0000004182; Apis mellifera.
CC InterPro: IPR003534; Royaljelly.
CC PANTHER: PTHR10009; Royaljelly; 1.
CC Pfam: PF03022; MRJP; 1.
CC PRINTS: PRO1366; ROYALJELLY.
CC Direct protein sequencing; Glycoprotein; Signal.
CC FT SIGNAL 1 17
CC CHAIN 18 452 Major royal jelly protein 2.
CC CARBOHYD 145 145 N-linked (GlcNAc...) (Potential).
CC CARBOHYD 178 178 N-linked (GlcNAc...) (Potential).
CC SEQUENCE 452 AA; 51074 MW; 25E5B621288FE189 CRC64;
CC -----
Query Match 58.58; Score 1634.5; DB 1; Length 452;
Best Local Similarity 68.88; Pred. No. 3.3e-95; Mismatches 72; Indels 11; Gaps 1;
Matches 305; Conservative 55;
QY 6 QRSANNLAKSMKVIYEWKHIDFDGSDERRDAAIKSGEFDHTKNYPFDVDRWRDKTFVT 65
DB 21 RENSPRNLEKSLNVIHEWKYFDYDFGSEERRQAAIQSGEYDHTKNYPFDVDRWRDKTFVT 80
QY 66 IERNNGVPSSLNVNKKGGPLLRYPYDWSFAKEDCGISVAFKIAVDKFDRLWVLD 125
DB 81 ILRYDGVPTLNVISGRTGGRLLKFPYDWSFAEFKDCSKIVSAFKIAIDKFDRLWVLD 140
QY 126 SGLVNNQPMCSPKLLTFDLKTSKLVQVEIPHNIAVNATTGMGELVSLAVQAIDRTNTM 185
DB 141 SGLVNRVTVPCAPKLHVFDLKTSHLQKQIEIPHNIAVNATTGMGELVSLAVQAIDRLANTL 200
QY 186 VYIADEKGEGLIMYQNSDDSFHRLTSNTFDYDPRYTKLTAVAGESFTVKNIGCGIALSPVT 245
DB 201 VYMAHKGDALIVYQNADDSFHRLTSNTFDYDPRYAKWTIDGESFTLKNIGCGIALSPVT 260
QY 246 NNLYSPLSHGLEYVDTEQFRNPQYENNVQEGSDIINTQSGFGKVSXKNGVLFGLV 305
DB 261 NNLYSPLSHGLEYVNTAPFMKSSQFGENNVOYQGSSEDIINTQSLAKAVSKNGVLFGLV 320
QY 306 GNSGIACVNEHQVLORESFDVVAQNETLQMIIVSMKIMENLPOSGRINDPEGNEYMLALS 365
DB 321 GNSAVGCWNEHQSLQRLQNEWVAQNDRTLQMIAGMKIKBELPHFVGSNGKPVKDEYMLVLS 380
QY 366 NRMQKIINDDFNFDVNRILGANVDDLMRNTCRGRYHNNQAGNQADNQADNQANNQ 425
DB 381 NRMQKIINDDFNFDVNRILGANVVKELIRNTHCVN-----NNQNDNIQNTNNQ 429
QY 426 NADQNANKQNGRQNDNRQNDN 448
DB 430 NDNNQNNKNNKNNQNNQNDN 452
RESULT 6
Q6W3E3_APIME
```

```
ID Q6W3E3_APIME PRELIMINARY; PRT; 437 AA.
AC Q6W3E3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Major royal jelly protein MRJP6.
OS Apis mellifera (Honeybee).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
OC Apidae; Apis.
OX NCBI_TaxID=7460;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15037093; DOI=10.1016/j.jinsphys.2003.09.008;
RA Albert S., Klaudiny J.;
RT "The MRJP/YELLOW protein family of Apis mellifera: identification of
RT new members in the EST library.";
RL J. Insect Physiol. 50:51-59(2004).
DR EMBL: AY13893; AAQ82184.1; -; mRNA.
DR Ensembl: ENSAPMG0000007331; Apis mellifera.
DR InterPro: IPR003534; Royaljelly.
DR Pfam: PF03022; MRJP; 1.
DR PRINTS: PRO1366; ROYALJELLY.
DR SEQUENCE 437 AA; 49786 MW; 1B654BB7A7F0B56D CRC64;
Query Match 56.28; Score 1572.5; DB 2; Length 437;
Best Local Similarity 72.28; Pred. No. 2.7e-91;
Matches 294; Conservative 56; Mismatches 56; Indels 1; Gaps 1;
QY 5 HQRSANNLAKSMKVIYEWKHIDFDGSDERRDAAIKSGEFDHTKNYPFDVDRWRDKTFV 64
DB 23 HQRSKSKNLEHSMNVVHEWKYIDYDFGSDERKQAAIQSGEYDHTKNYPFDVDRWRDKTFV 82
QY 65 TIERNNGVPSSLNVNKKGGPLLRYPYDWSFAKEDCGISVAFKIAVDKFDRLWVLD 124
DB 83 AVIRYDGVPTLNVIHEWKYIDYDFGSDERKQAAIQSGEYDHTKNYPFDVDRWRDKTFV 142
QY 125 DSGLVNNQPMCSPKLLTFDLKTSKLVQVEIPHNIAVNATTGMGELVSLAVQAIDRTNT 184
DB 143 DSGLVNNQPMCSPKLLTFDLKTSKLVQVEIPHNIAVNATTGMGELVSLAVQAIDRTNT 202
QY 185 MYIADEKGEGLIMYQNSDDSFHRLTSNTFDYDPRYTKLTAVAGESFTVKNIGCGIALSPV 244
DB 203 TVYIADRGDALIYQNSDDSFHRLTSNTFDYDPRYTKLTAVAGESFTVKNIGCGIALSPV 262
QY 245 TNLYSPLSHGLEYVDTEQFRNPQYENNVQEGSDIINTQSGFGKVSXKNGVLFGL 304
DB 263 TNLYSPLSHGLEYVNTAPFMKSSQFGENNVOYQGSSEDIINTQSLAKAVSKNGVLFGL 322
QY 305 VNSGIACVNEHQVLORESFDVVAQNETLQMIIVSMKIMENLPOSGRINDPEGNEYMLAL 364
DB 323 VNSAIGCWNEHQPLQRLQNEWVAQNDRTLQMIIVSMKIMENLPOSGRINDPEGNEYMLAL 382
QY 365 SNRMQKIINDDFNFDVNRILGANVDDLMRNTCRGRYHNNQAGNQNDN 411
DB 383 SNRMQKIINDDFNFDVNRILGANVNNLIKTRCAKSNQNDN-NNQN 428
RESULT 7
Q6IMJ9_APIME PRELIMINARY; PRT; 443 AA.
AC Q6IMJ9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Major royal jelly protein 7.
OS Apis mellifera (Honeybee).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
OC Apidae; Apis.
OX NCBI_TaxID=7460;
RN [1]
```

RP NUCLEOTIDE SEQUENCE.  
RX PubMed=15037093; DOI=10.1016/j.jinsphys.2003.09.008;  
RA Albert S., Klaudiny J.;  
RT "The MRJP/YELLOW protein family of Apis mellifera: identification of  
RL new members in the EST library.";  
RT J. Insect Physiol. 50:51-59(2004).  
CC -I- MISCELLANEOUS: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ third party annotation (TPA) entry.  
DR EMBL; BK001420; DAA01512.1; -; mRNA.  
DR Ensembl; ENSAPMG00000007331; Apis mellifera.  
DR InterPro; IPR003534; Royaljelly.  
DR Pfam; PF03022; MRJP; 1.  
DR PRINTS; PR01366; ROYALJELLY.  
SQ SEQUENCE 443 AA; 50541 MW; 8916272BA4DB421 CRC64;  
  
Query Match 53.2%; Score 1486.5; DB 2; Length 443;  
Best Local Similarity 66.9%; Pred. No. 7.4e-86;  
Matches 281; Conservative 53; Mismatches 73; Indels 13; Gaps 3;  
  
QY 6 QKSAANLAHSMKVIYEWKHIDFGSDERRDAAIKSGEDHTKNTYFPDVRDWDKTFVT 65  
DB 21 RENSARNLKNLSKVMHEWKYIDYDFGSEERKQAAIQSDEYDHTKNTYFPDVRDWDKTFVT 80  
  
QY 66 IERNNGVPSSLVNVTNKKGGPGLLRYPYDWSFAKYEDCSGIVSAFAKIAVDKFDRLWVLD 125  
DB 81 VLRYDGVPSLVNVTSEKNGRLLQYPYDWSWTKYKDCSGIVSAYSIAIDKFDRLWVLD 140  
  
QY 126 SGLVNNQPMCSKLLTFDLKTSKLVQVEIPHNIANNATTGMGELVSLVAQAIIDRTNTM 185  
DB 141 SGLVNNQPMCFKLLVFDLNSQLIKQVDIPHEIAVNTTTEQGRKLSLAQVAISSVNTL 200  
  
QY 186 VYIADEKGEGLIMYQNSDDSFHRLTSNTFYDPRYTKLTVAGESFTVKNGICGIALSPVT 245  
DB 201 VYIADNKGDLIVYQNSDDSFHRLTSNTFYDPRYTKMTEGESFTVQDGIYGMALSPMT 260  
  
QY 246 NNLYSPSSHGLVYVDTEQFRNPQYBENNVOYEGSQDILNTQSGFKVSKNGVLFGLV 305  
DB 261 NNLYSPPLASRLVYVNTKPIKSEYGENKQYGVQVDFNTQTAKAVSKNGILFGLV 320  
  
QY 306 GNSGIACVNEHQVLORESFDVVAQNEETLQIMVSKIMENLPQ-----SGRIINDPGNE 359  
DB 321 NNTAVGCVNEHQVLORENTDVAQNEETLQIMVGMKIKQLLPHIVIIDINIDN----E 376  
  
QY 360 YMLALSNEQMKIINNDNFNFRILGANVDDLMRNTCRGRYHQNAGNQADNADN 419  
DB 377 YMLVLTNRMQILNNDLNFNFRILGIGVSDLENTRC---TNFNTQNDSDSDNNDD 433  
  
RESULT 8  
Q6QN15 APICE PRELIMINARY; PRT; 579 AA.  
AC Q6QN15;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Major royal jelly protein MRJPs precursor.  
GN Name=MRJP5;  
OS Apis cerana (Indian honeybee).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;  
OC Apidae; Apis.  
OX NCBI\_TaxID=7461;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Hypopharyngeal gland;  
RA Sittipraneed S., Cenphakdee K.;  
RT "Molecular Cloning of Major Royal Jelly Protein (MRJP5) cDNA from Apis  
RL cerana in Thailand.";  
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY532369; AAS21320.1; -; mRNA.  
DR InterPro; IPR003534; Royaljelly.  
DR Pfam; PF03022; MRJP; 2.  
DR PRINTS; PR01366; ROYALJELLY.

KW Signal.  
FT SIGNAL 1 20 Potential.  
SQ SEQUENCE 579 AA; 68289 MW; 8218A56E120244F2 CRC64;  
  
Query Match 51.4%; Score 1437.5; DB 2; Length 579;  
Best Local Similarity 49.8%; Pred. No. 1.3e-82;  
Matches 280; Conservative 63; Mismatches 64; Indels 155; Gaps 3;  
  
QY 6 QKSAANLAHSMKVIYEWKHIDFGSDERRDAAIKSGEDHTKNTYFPDVRDWDKTFVT 65  
DB 24 RENSARNLANSMMVHWEKLYDVFSGDEKQAAIQSGEYDHTKNTYFPDVRDWDKTFVT 83  
  
QY 66 IERNNGVPSSLVNVTNKKGGPGLLRYPYDWSFAKYEDCSGIVSAFAKIAVDKFDRLWVLD 125  
DB 84 VLRYKGVPSLVNVTNKKGGPGLLRYPYDWSWANYKDCSGIVSAYKIAIDKFDRLWVLD 143  
  
QY 126 SGLVNNQPMCSKLLTFDLKTSKLVQVEIPHNIANNATTGMGELVSLVAQAIIDRTNTM 185  
DB 144 SGIIINTQPMCSKPLHVFDLNTSQIKQVMMPHDIAINATTGGGLENLVQAADPMTNL 203  
  
QY 186 VYIADEKGEGLIMYQNSDDSFHRLTSNTFYDPRYTKLTVAGESFTVKNGICGIALSPVT 245  
DB 204 VYIADNKGDLIVYQNSDDSFHRLTSNTFYDPRYTKMTEGESFTVQDGIYGMALSPMT 263  
  
QY 246 NNLYSPSSHGLVYVDTEQFRNPQYBENNVOYEGSQDILNTQSGFKVSKNGVLFGLV 305  
DB 264 NNLYSPPLASRLVYVNTKPIKSEYGENKQYGVQVDFNTQTAKAVSKNGVLFGLV 323  
  
QY 306 GNSGIACVNEHQVLORESFDVVAQNEETLQIMVSKIMENLPQSGRIIN----- 353  
DB 324 NNSAIGCVNEHQVLOREQNDVAQNEETLQTVVAMQKM-HLPQSNRMRMRMRMRMRMD 382  
  
QY 354 ----- 353  
DB 383 RMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMD 442  
  
QY 354 ----- 353  
DB 443 KMKMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMD 502  
  
QY 354 -----DPGEYMLALSNEQMKIINNDNFNFRILGANVDDLMRNT 397  
DB 503 MHRMGRMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMD 562  
  
QY 398 RCGRYHQNAGNQADNADN 419  
DB 563 RCANSNNQN-----DNQKH 578  
  
RESULT 9  
Q5VK56 APICC PRELIMINARY; PRT; 598 AA.  
AC Q5VK56;  
DT 01-FEB-2005 (TrEMBLrel. 29, Created)  
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)  
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)  
DE Major royal jelly protein MRJP5.  
OS Apis cerana (Oriental honeybee).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;  
OC Apidae; Apis.  
OX NCBI\_TaxID=94128;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Su S., Chen S., Albert S., Zhong B.;  
RT "Molecular cloning of MRJP5 cDNA from Apis cerana.";  
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY392757; AAR83082.1; -; mRNA.  
DR InterPro; IPR003534; Royaljelly.  
DR Pfam; PF03022; MRJP; 2.  
DR PRINTS; PR01366; ROYALJELLY.  
SQ SEQUENCE 598 AA; 70522 MW; DD931CA92F433DF2 CRC64;

Query Match	51.0%; Score 1427; DB 2; Length 598;
Best Local Similarity	48.0%; Pred. No. 6.4e-82;
Matches 279; Conservative	64; Mismatches 64; Indels 174; Gaps 3;
QY	6 QRKSANNLAHSMKVIYEWKHIDFSGDERRDAAIKSGFEDHTKNYPDPVDRWRDKTFVT 65
DB	24 RENSRNLANSNMVHBEWKVLDYDFGSDERQRAAIQSGEYDHTKNYPDPVDRWHDMTFVT 83
QY	66 IERNNGVPSSLVNVTNKKGGPPLLRYPPDWSFAKEDSCGIVSAFKIADVDFDLWLVD 125
DB	84 VLRYKGVPSLVNISKIGNGPPLLQPYPDWSMANYKDCSGIVSAFKIADKDFDLWLVD 143
QY	126 SGLVNNQPMCPKLLTFDLKTSKLVKQVEIPHNIANVATTGMLVSLAVQALDRNTM 185
DB	144 SGIINTQPMCPKLLHFDLNTSQIKQVMVPHDIAINATTGKGLLENLVQAMDPMNTL 203
QY	186 VYIADEKGEGLIMYQNSDDSFHRLTSTNTFDYDRTYKLTIVAGESFTVKGICGIALSPVT 245
DB	204 VYIADNKGDAIIVYQNSDDSFHRLTSTNTFDYDRTYKLTIVAGESFTVKGICGIALSPMT 263
QY	246 NNLVYSPSSHGLYVDTEQFRNPQYBENNVOYEGSDIILNTQSGKVSNGVLFGLV 305
DB	264 NNLVYSPASSLSLYIINTKPKMSQYGTNNVQHEGVQDIENFQSIKIMSKNGVLPGLM 323
QY	306 GNSGIACVNEHVLQRESFDVVAQNEHTLQMIYSKIMENLPQSGRIN----- 353
DB	324 NNSAIGCWNEHQPLQRQMDVAQNEHTLQTVVAMQKM-HLPQSNRMNRMHKMNVRNSMN 382
QY	354 ----- 353
DB	383 RMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMDRMD 442
QY	354 ----- 353
DB	443 RTNKMDRMDRMDIMDKMKNKMDRMDSMIRIDKMDRMDRMDRIDIMNRMDRMDTMDRID 502
QY	354 -----DEGNEYMLALSNRQKLIINDNF 378
DB	503 TMDRMDRMDRMDKMDKINKMHRMGRMDRMDRMDRMDRMDRMDRMDRMDRMDRMD 562
QY	379 NDVNFRIILGANVDDLMNTRCGRYVHNNQAGNQADNQNADN 419
DB	563 NEVNFRIILGANVDDLMNTRCGRYVHNNQAGNQADNQNADN-----DNQNKIN 597
RESULT 10	
MRJPI APIME	
ID	MRJPI APIME STANDARD; PRT; 432 AA.
AC	O18330;
DT	28-FEB-2003 (Rel. 41, Created)
DT	28-FEB-2003 (Rel. 41, Last sequence update)
DT	10-MAY-2005 (Rel. 47, Last annotation update)
DE	Major royal jelly protein 1 precursor (MRJP-1) (Bee-milk protein).
OS	Name=MRJPI;
GN	Apis mellifera (Honeybee).
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC	Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
OC	Apidae; Apis.
OX	NCBI_TaxID=7460;
RN	[1]
RP	NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 39-48; 63-71 AND 361-380.
RC	TISSUE=Hypopharyngeal gland;
RX	MEDLINE=98055707; PubMed=9395329;
RA	Ohashi K., Natori S., Kubo T.;
RT	"Change in the mode of gene expression of the hypopharyngeal gland
RT	cells with an age-dependent role change of the worker honeybee Apis
RT	mellifera L.";
RL	Eur. J. Biochem. 249:797-802 (1997).
RN	[2]
RP	NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 20-30.
RC	TISSUE=Head;
RX	MEDLINE=99007754; PubMed=9791542; DOI=10.1007/s000180050229;
RA	Schmitzova J., Klaudiny J., Albert S., Schroeder W., Schreckengost W.,

RA	Hanes J., Judova J., Simuth J.;
RT	"A family of major royal jelly proteins of the honeybee Apis mellifera
RL	Cell. Mol. Life Sci. 54:1020-1030(1998).
CC	-i- FUNCTION: May play an important role in honeybee nutrition. Most
CC	abundant protein found in the royal jelly which is the food of the
CC	queen honey bee larva. The royal jelly determines the development
CC	of the young larvae and is responsible for the high reproductive
CC	ability of the honeybee queen.
CC	-i- SUBCELLULAR LOCATION: Secreted.
CC	-i- TISSUE SPECIFICITY: Found in the hypopharyngeal glands of the
CC	worker honeybee.
CC	-i- DEVELOPMENTAL STAGE: Produced in the cephalic glands of both the
CC	nurse bee and the forager bee. This bee milk protein changes to
CC	alpha-glucosidase in accordance with the age-dependent role change
CC	of the worker bee.
CC	-i- SIMILARITY: Belongs to the major royal jelly protein family.
CC	-----
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use as long as its content is in no way modified and this statement is not
CC	removed.
CC	-----
DR	EMBL; D79207; BAA23639.1; -; mRNA.
DR	EMBL; AF000833; AAC61895.1; -; mRNA.
DR	Ensembl; ENSAPG00000007331; Apis mellifera.
DR	InterPro; IPR009007; Pept Aspartc_cat.
DR	InterPro; IPR003534; Royaljelly.
DR	PANTHER; PTHR10009; Royaljelly; 1.
DR	Pfam; PF03022; MRJP; 1.
DR	PRINTS; PR01366; ROYALJELLY.
DR	Direct protein sequencing; Glycoprotein; Signal.
KW	SIGNAL 1 19
FT	CHAIN 20 432 Major royal jelly protein 1.
FT	CARBOHYD 28 28 N-linked (GlcNAc...) (Potential).
FT	CARBOHYD 144 144 N-linked (GlcNAc...) (Potential).
FT	CARBOHYD 177 177 N-linked (GlcNAc...) (Potential).
SQ	SEQUENCE 432 AA; 48886 MW; 9F42BF08D34A1A7B CRC64;
Query Match	50.8%; Score 1420.5; DB 1; Length 432;
Best Local Similarity	65.3%; Pred. No. 1.1e-81;
Matches 265; Conservative	63; Mismatches 65; Indels 13; Gaps 2;
QY	12 NLAHSMKVIYEWKHIDFSGDERRDAAIKSGFEDHTKNYPDPVDRWRDKTFVTIERNG 71
DB	26 SLNKSPLILHEWKFFDYDFGSDERRQDAILSGEYDKNYPSIDIDQWHDKIFVTMLRYNG 85
QY	72 VPSSLVNVTNKKGGPPLLRYPPDWSFAKEDSCGIVSAFKIADVDFDLWLVDGLVNN 131
DB	86 VPSSLVNISKVGGDGGPPLLQPYPDWSFAKYDDCGSIVSASKLAIDKCDRLWLVDGLVNN 145
QY	132 NQPMCSKLLTFLDKTSKLVKQVEIPHNIANVATTGMLVSLAVQALDRNTM 188
DB	146 TQPMCSKLLTFLDTTSSQLLKQVEIPHVAVNATTGKRSLSSAVQSLDCNTSDTMVYI 205
QY	189 ADEKSGELIMYQNSDDSFHRLTSTNTFDYDRTYKLTIVAGESFTVKGICGIALSPVTNNL 248
DB	206 ADEKSGELIVYHNSDDSFHRLTSTNTFDYDRTYKLTIVAGESFTVKGICGIALSPVTNNL 265
QY	249 YVPSLSHGLYVDTEQFRNPQYBENNVOYEGSDIILNTQSGKVSNGVLFGLVNGS 308
DB	266 YVSPVASTSLYYNTQFRSTYDQNDIHVEGVQNLIDTQSSAKVSKVSGVLFGLVNGS 325
QY	309 GIACVNEHVLQRESFDVVAQNEHTLQMIYSKIMENLPQSGRINDEGNEYMLALSNRM 368
DB	326 ALGCWNEHRLTHERNIRTVAQSDETLQMIASMKIKEALPHVPIFDRIYINREYILVLSNKM 385
QY	369 QKIINDNFNDVNFRIILGANVDDLMNTRCGRYVHNNQAGNQADN 414
DB	386 QKWNNDNFDDVNFRIIMANVNELINTRC-----ENPDN 421



[illegible]



```
GN Name=MRJP4;
OS Apis cerana (Indian honeybee).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
OC Apidae; Apis.
OX NCBI_TaxID=7461;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Hypopharyngeal gland;
RA Sittipraneed S., Cenphakdee K.;
RT "Molecular Cloning of Major Royal Jelly Protein (MRJP4) cDNA from Apis
RT cerana in Thailand.";
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY323268; AAS21319.1; mRNA.
DR InterPro; IPR003534; Royaljelly.
DR Pfam; PF03022; MRJP; 1.
DR PRINTS; PR01366; ROYALJELLY.
KW Signal.
FT SIGNAL.
SQ SEQUENCE 1 20 Potential.
485 AA; 54970 MW; 278951523B9C4536 CRC64;

Query Match 50.1%; Score 1402; DB 2; Length 485;
Best Local Similarity 57.7%; Pred. No. 1.9e-80;
Matches 273; Conservative 82; Mismatches 106; Indels 12; Gaps 7;

Qy 2 AVNHQRKSANNLAHSMKVIYEWKHIDFDGSDERDDAAIKSGEPDHTKNYPFDVDRWRDK 61
Db | : | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 21 AVVRENSRKKLTNTLVNIHEWKTYVDYDFGSDKRAAIQSGYDRTKNYPFLDQWHDK 80
Db | : | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 62 TFWTIERNNGVPSSINVVTKKGGPLRPYPDWSFAKYEDCSGIVSAPKIAVDKFDRL 121
Db | : | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 81 TFWTMLRYDGVFPSSLVNVSDKTGNGGPLLQYPDWSFAKYEDCSGIVSANKIAIDYERL 140
Db | : | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 122 WVLDSGLVNNQPMCSPKLLTFDLKTSKLVKQVEIPHNIAYNATTGMGELVSLAVQAIDR 181
Db | : | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 141 WVLDSGLVNNIQPMCSPKLLAFDLTTSKLVKQVEIPHNDVAVNATTGGGLASLAVQAMDS 200
Db | : | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 182 TNTWVYIADEKGEGLIMYQNSDDSPHRLTSNTFDYDPRYTKLTVAGESFTVK---NGICG 238
Db | : | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 201 VNTWVYADNKKDALLVYQNADDSFHLSSHSI SNHNFSDKM--SQENLTILKEVDNRVFG 258
Db | : | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 239 IALSPVTNNLYYSPSSRHGLYYVDTEQFRNPQYENNVOYEGSODILNTQSFQKVVSKNG 298
Db | : | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 259 MALSSVTHNLYYSPSSPLSSQNLVYVNTSLMNSQNGNDVQYESVDVFSQLSAKAVSKNG 318
Db | : | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 299 VLFGLGVNGSIACVNEHQVLQRESFDVVAQNEETLQIVSMKIMENLPQSGRINDPEGN 358
Db | : | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 319 VLFPGFTWNT-LGCWNEHQSLDRQNDIVARN-ETLQWVVGMIKQNLPSQSKVNTQRN 376
Db | : | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 359 EYMLALSVMQKIINDDNFNDVNFRIILGANVDDLMRTTCGRYHYNQAGNQADNQAD 418
Db | : | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 377 EHLIALTNKKQDVLNNDLNLHVNFPQILDANVNDLIRNSRCANSDNQ---DNNQHNYN-H 432
Db | : | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 419 NQANNNQADNQKNGNQNDNRQNDKNGNQNDKNGNQNDKNGNQNDKNGNQNDKNG 471
Db | : | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 433 NQVRHSSKSDNQN-NNQHNQYHSSKSDNNDNNNNQAHSSKFDNQNNNQYN 484
Db | : | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

This Page Blank (upper)

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2006, 19:09:24 ; Search time 56.6049 Seconds  
(without alignments)  
765.341 Million cell updates/sec

Title: US-10-525-567-3

Perfect score: 2796

Sequence: 1 AAVNHQKSNANLAHSMKVI.....NNQNDNRNDNQVHSSKLH 524

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgm2\_6/ptodata/1/iaa/5 COMB.pep.\*  
2: /cgm2\_6/ptodata/1/iaa/6 COMB.pep.\*  
3: /cgm2\_6/ptodata/1/iaa/H COMB.pep.\*  
4: /cgm2\_6/ptodata/1/iaa/PTUS COMB.pep.\*  
5: /cgm2\_6/ptodata/1/iaa/RE COMB.pep.\*  
6: /cgm2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	247	8.8	1584	2	US-09-457-040B-27
2	232	8.3	162	2	US-09-270-767-33773
3	232	8.3	162	2	US-09-270-767-48990
4	220	7.9	954	2	US-09-570-856B-31
5	206	7.4	524	2	US-09-543-681A-4787
6	198.5	7.1	418	2	US-09-248-796A-20578
7	194.5	7.0	802	2	US-09-134-000C-5150
8	193	6.9	2391	1	US-08-446-855A-2
9	193	6.9	2391	2	US-09-150-741-2
10	182	6.5	961	2	US-09-487-558B-132
11	180	6.4	182	2	US-09-583-110-4334
12	180	6.4	186	2	US-09-107-433-3319
13	176	6.3	174	2	US-09-134-001C-4460
14	172.5	6.2	1162	1	US-08-728-323A-2
15	172.5	6.2	1162	2	US-09-298-568-2
16	172.5	6.2	1162	2	US-09-410-399-2
17	172.5	6.2	1162	2	US-09-894-273-2
18	169	6.0	666	2	US-09-487-558B-36
19	168.5	6.0	2184	2	US-09-417-485D-6
20	167.5	6.0	467	2	US-09-134-001C-3235
21	167.5	6.0	472	2	US-09-710-279-1902
22	166.5	6.0	696	2	US-09-248-796A-14133
23	163.5	5.8	110	1	US-08-569-166-34
24	163.5	5.8	133	2	US-09-710-279-1924
25	162.5	5.8	826	1	US-07-638-431-2
26	162.5	5.8	826	4	PCT-US92-00018-2
27	162	5.8	369	2	US-09-248-796A-23009

28	157.5	5.6	188	2	US-09-248-796A-21795	Sequence 21795, A
29	157.5	5.6	430	1	US-08-945-848-8	Sequence 8, Appli
30	156.5	5.6	213	2	US-09-248-796A-18858	Sequence 18858, A
31	155	5.5	263	2	US-09-248-796A-15305	Sequence 15305, A
32	154.5	5.5	1259	2	US-09-949-016-10366	Sequence 10366, A
33	148.5	5.3	758	2	US-09-487-558B-224	Sequence 224, App
34	147	5.3	124	2	US-09-248-796A-21892	Sequence 21892, A
35	146.5	5.2	596	2	US-09-752-165-2	Sequence 2, Appli
36	146.5	5.2	596	2	US-10-167-831-2	Sequence 2, Appli
37	145.5	5.2	602	2	US-09-248-796A-19204	Sequence 19204, A
38	145	5.2	3788	2	US-09-336-447A-76	Sequence 76, Appl
39	145	5.2	3788	2	US-09-952-267B-76	Sequence 76, Appl
40	144	5.2	1195	2	US-09-538-092-517	Sequence 517, App
41	143	5.1	588	2	US-09-248-796A-20839	Sequence 20839, A
42	143	5.1	657	2	US-09-248-796A-27210	Sequence 27210, A
43	143	5.1	1588	4	PCT-US93-07261-11	Sequence 11, Appl
44	143	5.1	1663	4	PCT-US93-07261-16	Sequence 16, Appl
45	142.5	5.1	78	2	US-09-328-352-6915	Sequence 6915, Ap

## ALIGNMENTS

## RESULT 1

US-09-457-040B-27  
; Sequence 27, Application US/09457040B  
; Patent No. 6387641  
; GENERAL INFORMATION:  
; APPLICANT: Vertex Pharmaceuticals Incorporated  
; APPLICANT: Bellon, Steve  
; TITLE OF INVENTION: Crystallized P38 Complexes  
; FILE REFERENCE: VPI/98-14  
; CURRENT APPLICATION NUMBER: US/09/457,040B  
; CURRENT FILING DATE: 1999-12-08  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 27  
; LENGTH: 1584  
; TYPE: PRT  
; ORGANISM: DICDI - Dictyostelium Discoideum  
US-09-457-040B-27

Query Match 8.8%; Score 247; DB 2; Length 1584;  
Best Local Similarity 24.7%; Pred. No. 1.5e-14;  
Matches 74; Conservative 30; Mismatches 108; Indels 88; Gaps 9;

Qy	255	SHGLYYVDTFQFRNPQYEENNVQEGSQDILNTQSGKVKVSKGVFLGLVGN-	-----	308
Db	295	SYGYHNDGGRKFR---WREBPGVNEG-----BSYGSSYKKGDIIGGLSFTSREIFFT	344	
Qy	309	-----GIACVNEHQVLORESFDVVAQNEETLQMIVSMKIMENLPQSGRINDPEGNYML	362	
Db	345	KNGMYLTAFSNVYGVF-----YPSVAFNEPGIS-	-----	381
Qy	363	ALSNRMQKIIINDNFNDVNFRIILGANVDDILMRNTRCGRVHNQAGNQADNQADNQNA	422	
Db	382	-----PKFSQVTLMLKNVNSTIL-----VPGNNNNNNNNNNNNNN	419	
Qy	423	N-----NQADNQNAKQNGKQNDNRQNDKQNGKQNDKQNGKQNDKQNGKQNDKQNG	462	
Db	420	NIIGNGKITTTTSTSPSSINNNDISSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	479	
Qy	463	NQNDKQNGKQNGKQNDKQNGKQNDKQNGKQNDKQNGKQNDKQNGKQNDKQNGKQNDKQNG	522	
Db	480	NSSNTNNNNNTTT-NNNNNSNNNNNNNNNSNSNSNNNNNNNNNNNNNNNNNNNNNN	538	

## RESULT 2

US-09-270-767-33773  
; Sequence 33773, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.

```
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 33773
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-33773

Query Match
Best Local Similarity 8.3%; Score 232; DB 2; Length 162;
Matches 52; Conservative 25; Mismatches 43; Indels 18; Gaps 5;

Qy 44 EFQHTKYPDFDVRDRDKTFTVTTIER--NGVPSSINVT--NKKGKGGPLLRPYDPWSF--- 98
Db 7 DFVPKNLPLGIDVHNNRLFTVTPRWKNGVPASLGTLPPFSKGVKGPKALPKYPNWEAHGN 66

Qy 99 AKYEDCSGIVSAFKIAVDKFDRLWLVDGLVN---NNQPMCSPKLLTFDLKTSK----- 150
Db 67 PNPDCSKLSVYRTAVDRCDRIWLIDSGIVNATINLQICPPKIVVYDLKSDDELIVRYN 126

Qy 151 -----VKQVEIPHNIAYN 163
Db 127 LEASHVKQDSLHSGNIVVD 144

RESULT 3
US-09-270-767-48990
; Sequence 48990, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48990
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-48990

Query Match
Best Local Similarity 8.3%; Score 232; DB 2; Length 162;
Matches 52; Conservative 25; Mismatches 43; Indels 18; Gaps 5;

Qy 44 EFQHTKYPDFDVRDRDKTFTVTTIER--NGVPSSINVT--NKKGKGGPLLRPYDPWSF--- 98
Db 7 DFVPKNLPLGIDVHNNRLFTVTPRWKNGVPASLGTLPPFSKGVKGPKALPKYPNWEAHGN 66

Qy 99 AKYEDCSGIVSAFKIAVDKFDRLWLVDGLVN---NNQPMCSPKLLTFDLKTSK----- 150
Db 67 PNPDCSKLSVYRTAVDRCDRIWLIDSGIVNATINLQICPPKIVVYDLKSDDELIVRYN 126

Qy 151 -----VKQVEIPHNIAYN 163
Db 127 LEASHVKQDSLHSGNIVVD 144

RESULT 4
US-09-570-856B-31
; Sequence 31, Application US/09570856B
; Patent No. 6682923
; GENERAL INFORMATION:
; APPLICANT: Bentzien, Joerg M
; APPLICANT: Dahiyat, Basil I
; TITLE OF INVENTION: NOVEL THERMOSTABLE ALKALIPHILIC XYLANASE
; FILE REFERENCE: A-67478-1/RPT/RMS/RMK
```

```
; CURRENT APPLICATION NUMBER: US/09/570,856B
; CURRENT FILING DATE: 2002-04-15
; PRIOR APPLICATION NUMBER: US 60/133,714
; PRIOR FILING DATE: 1999-05-12
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 954
; TYPE: PRT
; ORGANISM: Ruminococcus flavefaciens
US-09-570-856B-31

Query Match
Best Local Similarity 7.9%; Score 220; DB 2; Length 954;
Matches 67; Conservative 26; Mismatches 44; Indels 50; Gaps 12;

Qy 365 SNRMQKIIINDEN--FNDVNERILGANVDDLMRETRCGRYHONAGNONAD-----NQNA 417
Db 415 NNNQOO--NNDMNQNNQNNW-----NNQQNNNDWQNNQ--GQQNNNDWQNNQNN 464

Qy 418 DNQANNNQNA-----DNQANQKQNGRQNDNRQNDNK---QNGNRQNDNRQNDNK----- 461
Db 465 WQNNNNQQAAGWDNNNNWQWDQNNQNNQNNQNNQNNQNNQNNQNNQNNQNNQNNQNN 524

Qy 462 -GNRQND-----NKQNGNRQNG---NKQNDKQNGNRQND-----DNKRGNR---QNDNQ 504
Db 525 QGQQNNNDWQNNQNNQNNQNGWDNNQNNQNNQNNQNNQNNQNNQNNQNNQNNQNNQNDW 584

Qy 505 NNQNDNN 511
Db 585 NNQQQNN 591

RESULT 5
US-09-543-681A-4787
; Sequence 4787, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709,1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4787
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4787

Query Match
Best Local Similarity 7.4%; Score 206; DB 2; Length 524;
Matches 65; Conservative 23; Mismatches 74; Indels 26; Gaps 10;

Qy 338 VSMKIMENLPQSGRINDPEGNEYMLALSNMQK-----IINNDFNFNDVNFRIILGANVDDL 393
Db 141 IALKCLYEKLDKNEELNDQIFAEIYIELDNKIRSGEIDLKNSFNQVNFSAIRQLVSKA 200

Qy 394 MRNTRCGRYHONAGNONADNQANQNNQNNQNNQNNQNNQNNQNNQNNQNNQNNQNNQNN 453
Db 201 IEN-----DNTWLDTEN-DNINLENDNINLEN-DNINLENDNINLENDNINLEN-DNIN 251

Qy 454 RQNDN-----KQNGNRQNDKQNGRQNDKQNGRQNDKQNGRQNDKQNGRQNDKQNGRQND 509
Db 252 LENDNINLENDNINLENDN--VNLENDNVLENDNV---NLENDNV---NLENDNVLEND 304

Qy 510 --NNRNDN 515
Db 305 NINLENDN 312
```

**RESULT 6**

US-09-248-796A-20578  
; Sequence 20578, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstein et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS  
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; PRIOR FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 20578  
; LENGTH: 418  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-09-248-796A-20578

Query Match                7.1%; Score 198.5; DB 2; Length 418;  
Best Local Similarity     25.1%; Pred. No. 1e-10;  
Matches      67; Conservative     53; Mismatches    98; Indels     49; Gaps     12;

QY          268 NPQEENNQQVEGSDILNTOSFGKGVSKNGVLFGLVGNSGIACVNEHQVLQRESFDVV 327  
DB           |::|:|:|:|:~::~||:~::~||:~::~||:~::~||:~::~||:~::~||:~::~||:~::~||  
QY          328 AQNBETLQMIVSMKMENLPQSGRINDEPGEYMALSNRMOKI-----INNND 375  
DB           ||::|:~::~||:~::~||:~::~||:~::~||:~::~||:~::~||:~::~||:~::~||  
QB          224 KRNLSLIKELPPPI-ILENVP--SSLPKKQPQNLLNIIPKARSKFTPTTTPRQFTHNNDD 281  
QY          376 FNFNVDVFRIIGANVDDLWRNYRCRGYHNQAGNQADNOADNONANNONADNONANKQ 435  
DB           |::|:~::~||:~::~||:~::~||:~::~||:~::~||:~::~||:~::~||:~::~||  
QB          282 F---EVTHWGNKN-----NINHHRRNNVSDFSRSNDOHRRNYSDFSRDHQG 326  
QY          436 NGNRQMDNRQDNKDQNGNRQDNDKQNGNRQDND--KQNGNRQDKQNKQNGNRQDND 493  
DB           ::||:~::~||:~::~||:~::~||:~::~||:~::~||:~::~||:~::~||:~::~||  
QB          327 --HRQPQQNNYYNNNNNNNNNNNNNNNNNNNNNGFKVRSGSVRYNVNYYNNN--NGNANG 382  
  
QY          494 KRNGY----QRNDQNMQRNDNRNDQOV 517  
DB           ::||:~::~||:~::~||:~::~||:~::~||:~::~||:~::~||:~::~||:~::~||  
QB          383 NNHGKPKFRSTYNNNNNNNNNNNNNNKRNSQL 409

**RESULT 7**

US-09-134-000C-5150  
; Sequence 5150, Application US/09134000C  
; Patent No. 6617156  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 032796-032  
; CURRENT APPLICATION NUMBER: US/09/134,000C  
; PRIOR FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/055,778  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 6812  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 5150  
; LENGTH: 802  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-134-000C-5150

Query Match                7.0%; Score 194.5; DB 2; Length 802;  
Best Local Similarity     38.3%; Pred. No. 7.4e-10;  
Matches      49; Conservative    13; Mismatches    45; Indels     21; Gaps     6;

```
Db 1277 NH-----LYNEVDDKDTQLHKNNNNNM-----NSGVNCKLKNESYG 1318
Qy 218 PRYKLTVAGESFTVKGIC-GIALS---PVTNNLYSPLS-----SH 256
Db 1319 YNNSNCINTNNIENNICHDISINKNIKVTIINNNSNISNNENVTNLNCVSRAGSH 1378
Qy 257 GLYVDTQFRPNQYENNVQYEGSQD--ILNTQSFQKVKVSKGVFLGLVGNSGIACVN 314
Db 1379 HIY-----GKEEKSI---GSDDTNILSAQN-----SNNFSCNN 1409
Qy 315 EH-----QVLO-----RESFDVVAQNEETLQIMVSKIMENLPQSGRINDPEGNEYM 361
Db 1410 ENMNKANVDVNVLENDTKKREDINTTTVFMEGQNSVINNNKEN-----SLLKGEEDIVM 1466
Qy 362 LAL--SNRMQKIINN-----DFNFNDVN-----FRILGAN---VDDLNRNT--- 397
Db 1467 VNLKKNYNSVINNVDCRKKDQKNDCKTYKKNYKDKMGLNINIVDELSNGTSHS 1526
Qy 398 -----RCGRYHNQVAGN-----Q 410
Db 1527 TNDHLYLDNFNTSDEBTGNKNDMYLSKEKSIENKPNGSYYVVDVSVYNNYKINKMKE 1586
Qy 411 NADQNADQNANQNADQNANKQ-----NGNRQNDNRND-----NKONGN 453
Db 1587 LIDNENLNDYNNNVNMCNSYNNNASAFVNGKDRNDNLENDCEKMDHTYKHYNRLN-N 1645
Qy 454 RQNDNKQKQNGRQNG-RNQKQNGRQNDKQNGRQNGRQNDQNN 506
Db 1646 RRTSTNERMMLVMVNEKESNEHKGHRRLNGLNKKNEKMKKGNKDKKNYHYVNHKRN 1705
Qy 507 QNDNRNDNQVHH 519
Db 1706 EYNSNNIESKFNN 1718
```

## RESULT 9

```
US-09-150-741-2
; Sequence 2, Application US/09150741
; Patent No. 6183996
; GENERAL INFORMATION:
; APPLICANT: Stewart et al.
; TITLE OF INVENTION: Nucleotide Sequence Encoding Carbanoyl Phosphate
; Patent No. 6183996
; TITLE OF INVENTION: Synthesase II
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/150,741
; CURRENT FILING DATE: 1998-09-10
; EARLIER APPLICATION NUMBER: PL6380
; EARLIER FILING DATE: 1992-12-16
; EARLIER APPLICATION NUMBER: AU93/00617
; EARLIER FILING DATE: 1993-12-02
; EARLIER APPLICATION NUMBER: 08/446,855
; EARLIER FILING DATE: 1995-07-06
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2391
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: protein
US-09-150-741-2
```

```
Query Match 6.9%; Score 193; DB 2; Length 2391;
Best Local Similarity 19.8%; Pred No. 6.4e-09;
Matches 133; Conservative 106; Mismatches 204; Indels 230; Gaps 33;

Qy 5 HORKSANNLAHSMKVIYEWKHIDPDF-----GSDERRDAAIKSGFDHTKNYFQDV 55
Db 1118 HQ---AFLNMPDKIHETHIDWFLHFKFYNIYNLQNKLTLEQLSPNDLKY--FKK 1172
Qy 56 DRWRDK-----TPVTIERNNGVP--SSLNVVTVNKKGGGFLLRPPYDPWDSFAKYEDCSGI 107
```

```
Db 1173 HGFSDKQIAHYLSFNTSDNNNNNNNISCRTVEN-----DVMKYREKGL 1217
Qy 108 VSAFKIADVDF-----DRLWLDGSLVNNNQ--MCSPKLLTPDLTKYSLKLVQVEIP 157
Db 1218 FPHIKV-IDTLSAEPALTNLYLYTYQGQHDVPLPLNMRKKICTLNNKRNANKKHVHK 1276
Qy 158 HNIAVNATTMGELYSVLAQADRTNTWYIADKEGEGLIYQNSDDSFRLTSTNFTDYD 217
Db 1277 NH-----LYNEVDDKDTQLHKNNNNNM-----NSGVNCKLKNESYG 1318
Qy 218 PRYKLTVAGESFTVKGIC-GIALS---PVTNNLYSPLS-----SH 256
Db 1319 YNNSNCINTNNIENNICHDISINKNIKVTIINNNSNISNNENVTNLNCVSRAGSH 1378
Qy 257 GLYVDTQFRPNQYENNVQYEGSQD--ILNTQSFQKVKVSKGVFLGLVGNSGIACVN 314
Db 1379 HIY-----GKEEKSI---GSDDTNILSAQN-----SNNFSCNN 1409
Qy 315 EH-----QVLO-----RESFDVVAQNEETLQIMVSKIMENLPQSGRINDPEGNEYM 361
Db 1410 ENMNKANVDVNVLENDTKKREDINTTTVFMEGQNSVINNNKEN-----SLLKGEEDIVM 1466
Qy 362 LAL--SNRMQKIINN-----DFNFNDVN-----FRILGAN---VDDLNRNT--- 397
Db 1467 VNLKKNYNSVINNVDCRKKDQKNDCKTYKKNYKDKMGLNINIVDELSNGTSHS 1526
Qy 398 -----RCGRYHNQVAGN-----Q 410
Db 1527 TNDHLYLDNFNTSDEBTGNKNDMYLSKEKSIENKPNGSYYVVDVSVYNNYKINKMKE 1586
Qy 411 NADQNADQNANQNADQNANKQ-----NGNRQNDNRND-----NKONGN 453
Db 1587 LIDNENLNDYNNNVNMCNSYNNNASAFVNGKDRNDNLENDCEKMDHTYKHYNRLN-N 1645
Qy 454 RQNDNKQKQNGRQNG-RNQKQNGRQNDKQNGRQNGRQNDQNN 506
Db 1646 RRTSTNERMMLVMVNEKESNEHKGHRRLNGLNKKNEKMKKGNKDKKNYHYVNHKRN 1705
Qy 507 QNDNRNDNQVHH 519
Db 1706 EYNSNNIESKFNN 1718

RESULT 10
US-09-487-558B-132
; Sequence 132, Application US/09487558B
; Patent No. 6949356
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. 6949356man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; FILE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.130
; CURRENT APPLICATION NUMBER: US/09/487,558B
; CURRENT FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/487,558
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 446
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 132
; LENGTH: 961
; TYPE: PRT
```



```

THERAPEUTICS
NUMBER OF SEQUENCES: 5206
CORRESPONDENCE ADDRESS:
ADDRESS: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/ 085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 3319:
SEQUENCE CHARACTERISTICS:
LENGTH: 186 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...186
SEQUENCE DESCRIPTION: SEQ ID NO: 3319:
US-09-107-433-3319
Query Match 6.4%; Score 180; DB 2; Length 186;
Best Local Similarity 33.3%; Pred. No. 1.8e-09;
Matches 54; Conservative 27; Mismatches 49; Indels 32; Gaps 10
QY 393 LMENTRCGRVHNAG---NQAD--NQADNQVAN--NQADNQN-----ANKONGNRQN 441
DB 14 LLQNCVQNVQLNLRSKWNQNLRLNRKNQNLRLNRSKNQNLRLNRSKNQNLRLNR 73
QY 442 DNFRQNDNKQ-NGNRQNDNKQ-NGNRQNDN-----KONGNRQNGKQNDKQ-NGNRQN 491
DB 74 RSRKNQNLRLNRSKNQNLRLNRSKNQNLRLNRSKNQNLRLNRSKNQNLRLNRSKN 133
QY 492 DNKR-----NGNRQNDNQND-----NNRNDNQVHHSSK 522
DB 134 QNOLRLNRSKNQNLRLNRKNQNLRLNRKNQNLRLNRKNQNLRLNRSHK 175
RESULT 13
US-09-134-001C-4460
Sequence 4460, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHY
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964

```

Query Match 6.2%; Score 172.5; DB 1; Length 1162;



Db 142 IGWTTNCPYAVNVFLDTTTRIRRYELP-GVDTNPNTFIAMI-----AVDIGKNCD 194  
QY 184 TMVYIADEKEGLIMYQNSDDSFHRLTSNTFDY-DPRYTKLTVAGSFYVKNIGICGIAL 241  
Db 195 AYAFADDELGYGLIAYSWELNKSWSAHSYFFPDPLRGDPNVAGINFQWGESIGFMSL 254  
QY 242 SPVNTN----LYYSPSSHGLYYVDTEQFRNPQVEENNVOYEGSQDIL-----NTQSF 290  
Db 255 SPIRSDGYRTLYRSPSLASHRQFAVSTRILRDETREDSY-----HDFVALDERGPNSHTT 309  
QY 291 GKUVSKNGVLFLGVNGSGTACVNEHQVLQRESFDVVAQNEETLQIMVSKIMENLPQS 350  
Db 310 SRVMSDDGIELFLNIDQVAGCVHSSMPYSPQFHGIVDRDDVGLVFPADVKIDE----- 363  
QY 351 RINDPEGNEYMALSNRMQKIINDFNFNDVNFRIILGANVDDLMNRNTRCGRYHN 404  
Db 364 -----NKNVWVLSDRMPVFLSLDDYSNTNFIYTAPLATIENTVCDLRNN 410

## RESULT 2

US-11-097-143-25233  
; Sequence 25233, Application US/11097143  
; Publication No. US20050208558A1  
; GENERAL INFORMATION:  
; APPLICANT: Venter, J. Craig  
; APPLICANT: et al.  
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
; TITLE OF INVENTION: DROSOPHILA GENES.  
; FILE REFERENCE: CL000728  
; CURRENT APPLICATION NUMBER: US/11/097,143  
; CURRENT FILING DATE: 2005-04-04  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR FILING DATE: 1999-10-28  
; PRIOR FILING DATE: 1999-10-28  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR FILING DATE: 2000-01-12  
; PRIOR FILING DATE: 2000-01-12  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR FILING DATE: 2000-03-23  
; NUMBER OF SEQ ID NOS: 43008  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 25233  
; LENGTH: 426  
; TYPE: PRT  
; ORGANISM: DROSOPHILA  
US-11-097-143-25233

Query Match 15.1%; Score 422.5; DB 6; Length 426;  
Best Local Similarity 26.2%; Pred. No. 5.4e-26;  
Matches 111; Conservative 83; Mismatches 155; Indels 75; Gaps 13;  
QY 19 VIYEWKHIDPFGSDERRDAAIKSGDFDHTKNTYFPDVRW-----RDKTFVFTIER-NNG 71  
Db 34 IVFEWKNLQYFFSEQRDQVLRNGRYNPDSPIDIDVYPPNGGPPRHFTVSPRFGQ 93  
QY 72 VPSLNVVTKKGGKGLLPRLPY--DWSFAYEDCSGVSAPKIAVDKPRDLVLSGLV 129  
Db 94 VPFSGLVTVTVQRENGSEIOAYFSYQWHSHGANCGLTSVYRVHIDACQMWVLSGEI 153  
QY 130 NNNQPMCSKLLTFLDKTSKLQVKEIPHNIATNATGMGELVSLVAQIAIDR-----TN 183  
Db 154 EFVQ-HCAPQVMVFDLATDQLIHRYLPE---TSYKAKVSFRVNIPIADIRPPPSGCKD 209

QY 184 TMVYIADEKEGLIMYQNSDDSFHRLTSNTFDY-DPRYTKLTVAGSFYVKNIGICGIALS 242  
Db 210 VFAYLADPTSKAIVVVDVVGQSSWRI-ENKFTYPDAKFGHTTVAGESFELLDGFLALATT 268  
QY 243 P-----VTNNLYYSPSSHGLYYVDTEQFRNPQVEENNVOYEGSQDILNT----- 287  
Db 269 PLGLGURHLIFALS-----NELELAIPDLINNATNKHQKGLSSS 309  
QY 288 -----QSPGKVVSKNGVLFLGVNGSGTACVNEHQVLQRESFDVVAQNEETLQ 335  
Db 310 LSEFTVLGKGIQASHAISROGFLFCGFLPIGIFGWDIRRPYRNVKLLAINPATLQ 369  
QY 336 MIVSKIMENLPQSGRINDPEGNEYMALSNRMQKIINDFNFNDVNFRIILGANVDDLMR 395  
Db 370 FVSGMKIVRR-PADGR-----EELWLLSLDRLOKIFAGTIDYREINRYVMRCDVDDLLQ 421  
QY 396 NTRC 399  
Db 422 GRCC 425

## RESULT 3

US-11-097-143-40164  
; Sequence 40164, Application US/11097143  
; Publication No. US20050208558A1  
; GENERAL INFORMATION:  
; APPLICANT: Venter, J. Craig  
; APPLICANT: et al.  
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
; TITLE OF INVENTION: DROSOPHILA GENES.  
; FILE REFERENCE: CL000728  
; CURRENT APPLICATION NUMBER: US/11/097,143  
; CURRENT FILING DATE: 2005-04-04  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR FILING DATE: 1999-10-28  
; PRIOR FILING DATE: 1999-10-28  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR FILING DATE: 2000-01-12  
; PRIOR FILING DATE: 2000-01-12  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR FILING DATE: 2000-03-23  
; NUMBER OF SEQ ID NOS: 43008  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 40164  
; LENGTH: 453  
; TYPE: PRT  
; ORGANISM: DROSOPHILA  
US-11-097-143-40164

Query Match 15.0%; Score 418.5; DB 6; Length 453;  
Best Local Similarity 27.9%; Pred. No. 1.3e-25;  
Matches 122; Conservative 84; Mismatches 194; Indels 37; Gaps 12;  
QY 16 SMKVIYEWKHIDPFGSDERRDAAIKSGDFDHTKNTYFPDVRWRDKTFVFTIER-NNGVPS 74  
Db 21 NLRVAYEWMDFKYPANPDQWSAIEERGEFKPANVIPFGLVAGHRLFVTLPRWRDGVPA 80  
QY 75 SLNVV-TNKKGGKGLLPRLPYDPDWSFAYEDCS-GIYSAPKIAVDKPRDLVLD---SGLV 129  
Db 81 SLAYLDLNDTSSKGPALKPFPSPQAHNQAEPPELVSPFRVADRCRCRLVDSRISGLV 140  
QY 130 NNNQPMCSKLLTFLDKTSKLQVKEIPHNIATNATGMGELVSLVAQIAIDRNTVYIA 189  
Db 141 EQTKIYGAQALLVYDLHNDLLRRHVLPAQOLKQGSIL-----LANLAVEDSDCENTPAYAA 196





Qy 234 NGICGIALSPV---TNNLYSPSSHGLYYVDTEQPRNPQYBENNVOYEGSOD--ILNT 287  
 Db 267 DGIFSATLSGYKPGDKSDVFFHPMASTNEFVSN---RVLQQEFNAARSDHGDDFHLLGT 323  
 Qy 288 Q-----SFGKWSKNGVLFLGLVNGSGIACVNEHQVLORESFDVVAQNEBTLQMIYSMK 341  
 Db 324 RGPSTQTMKYPDRTGVIFFAEVQKSGVGCKWTSKPFSTENHGSVYSNSS----- 374  
 Qy 342 IMENLPOSGRINDPEGNEYMLALSNNRMQKIINDNFNDVNFRLIGANVDDLMNRTRC 399  
 Db 375 --EMIVPSDLTIDBEG--YIWMNSNMPFIYVSKLDVEKYNFRIWQSTLLAKRGTV 428

RESULT 8  
 US-11-097-143-25236  
 ; Sequence 25236, Application US/11097143  
 ; Publication No. US20050208558A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Venter, J. Craig  
 ; APPLICANT: et al.  
 ; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
 ; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
 ; TITLE OF INVENTION: DROSOPHILA GENES.  
 ; FILE REFERENCE: CL000728  
 ; CURRENT APPLICATION NUMBER: US/11/097,143  
 ; CURRENT FILING DATE: 2005-04-04  
 ; PRIOR APPLICATION NUMBER: 60/157,832  
 ; PRIOR FILING DATE: 1999-10-05  
 ; PRIOR APPLICATION NUMBER: 60/160,191  
 ; PRIOR FILING DATE: 1999-10-19  
 ; PRIOR APPLICATION NUMBER: 60/161,932  
 ; PRIOR FILING DATE: 1999-10-28  
 ; PRIOR APPLICATION NUMBER: 60/164,769  
 ; PRIOR FILING DATE: 1999-11-12  
 ; PRIOR APPLICATION NUMBER: 60/173,383  
 ; PRIOR FILING DATE: 1999-12-28  
 ; PRIOR APPLICATION NUMBER: 60/175,693  
 ; PRIOR FILING DATE: 2000-01-12  
 ; PRIOR APPLICATION NUMBER: 60/184,831  
 ; PRIOR FILING DATE: 2000-02-24  
 ; PRIOR APPLICATION NUMBER: 60/191,637  
 ; PRIOR FILING DATE: 2000-03-23  
 ; NUMBER OF SEQ ID NOS: 43008  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 25236  
 ; LENGTH: 409  
 ; TYPE: PRT  
 ; ORGANISM: DROSOPHILA  
 US-11-097-143-25236

Query Match 11.4%; Score 319.5; DB 6; Length 409;  
 Best Local Similarity 27.3%; Pred. No. 1.5e-17;  
 Matches 112; Conservative 70; Mismatches 167; Indels 61; Gaps 18;  
 Qy 17 MKVYEWKHIDFGSDERRDAALKSGSFDHTKYNPFDV-----DRWRDRTFTVIER-NN 70  
 Db 22 LHTLHOWTNLSIG-----DLSKGNRF-----LPVDVDIEYGDGHRHRTFTIPRLGM 69  
 Qy 71 GVPSSLVNVTNKKGK--GGPLLRPVP--DWSFAKYDCSGIVSAFKIAVDKFDRLWLVDLS 126  
 Db 70 ATPFTLATVIAHNELVENPKLEYPNEWHVPP--NNGSGITSARTIYIDECRWLVVDS 128  
 Qy 127 GLVNNQPMCSPKLLTFLDKTSKLVQVEIAPHNIAVNATTGMLGSLVLAQVADR----- 181  
 Db 129 GGVNSLQ-LCPPIQLTDLVQKDELQVHALPPDSYI---PSVSIPTALVVDLAERGTPNR 184  
 Qy 182 -TNTMVIYADEKGLIMYQNSDDSFHRLTNTFDYDPRYKLTIVAGESFTVKGICGIA 240  
 Db 185 CVGGRAYIADAGYGLIVFDSITGRSWRIEHSMKPSP-----LLRLGRSSNSQAGIFTVS 240  
 Qy 241 LSP---VTNNLYSPSSHG-----LYYVDTEQF--RNQYBENNVOYEGSODILNTQSGF 291

Db 241 LSPSEVEDRPLYFHTLNSFNEMRPVLSLINNETFWKSANASRDSFHSGLTRGI---OCES 297  
 Qy 292 KVVSKNGVLFLGLVNGSGIACVNEHQVLORESFD--VVAQNEETLQMIYSMKIMENLPQS 349  
 Db 298 EVMQSGNLYCSLI--SIGALVKEESVSNYTTADDLRVAVNPHKIRFVTGLKINRN----- 352  
 Qy 350 GRINDPEGNEYMLALSNNRMQKIINDNFNDVNFRLIGANVDDLMNRTRC 399  
 Db 353 -----SKGBEELWALSQPKLFVGGDLPAANEVKFQIIGCRTADILLANTPC 397

RESULT 9  
 US-11-097-143-41712  
 ; Sequence 41712, Application US/11097143  
 ; Publication No. US20050208558A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Venter, J. Craig  
 ; APPLICANT: et al.  
 ; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
 ; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
 ; TITLE OF INVENTION: DROSOPHILA GENES.  
 ; FILE REFERENCE: CL000728  
 ; CURRENT APPLICATION NUMBER: US/11/097,143  
 ; CURRENT FILING DATE: 2005-04-04  
 ; PRIOR APPLICATION NUMBER: 60/157,832  
 ; PRIOR FILING DATE: 1999-10-05  
 ; PRIOR APPLICATION NUMBER: 60/160,191  
 ; PRIOR FILING DATE: 1999-10-19  
 ; PRIOR APPLICATION NUMBER: 60/161,932  
 ; PRIOR FILING DATE: 1999-10-28  
 ; PRIOR APPLICATION NUMBER: 60/164,769  
 ; PRIOR FILING DATE: 1999-11-12  
 ; PRIOR APPLICATION NUMBER: 60/173,383  
 ; PRIOR FILING DATE: 1999-12-28  
 ; PRIOR APPLICATION NUMBER: 60/175,693  
 ; PRIOR FILING DATE: 2000-01-12  
 ; PRIOR APPLICATION NUMBER: 60/184,831  
 ; PRIOR FILING DATE: 2000-02-24  
 ; PRIOR APPLICATION NUMBER: 60/191,637  
 ; PRIOR FILING DATE: 2000-03-23  
 ; NUMBER OF SEQ ID NOS: 43008  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 41712  
 ; LENGTH: 379  
 ; TYPE: PRT  
 ; ORGANISM: DROSOPHILA  
 US-11-097-143-41712

Query Match 10.8%; Score 302.5; DB 6; Length 379;  
 Best Local Similarity 24.7%; Pred. No. 3.3e-16;  
 Matches 95; Conservative 70; Mismatches 181; Indels 39; Gaps 11;  
 Qy 35 RRDAAIKSGFDHTKYNPFDVDRWRDRTFTVIERNN--GVPSLVNV---TNKKGKGGPLL 91  
 Db 13 RADSDPDNASYIPYNNVPMGATHEPRGLFVTMPRRRVGIPSTLNYIDLAEDGSNNRSPKL 72  
 Qy 92 PYPDWSFAKYE--DCSGIVSAFKIAVDKFDRLWLVDGLV---NNQPMCSPKLLTFLDKLT 147  
 Db 73 AYPFALNQNFAENLVSVYRTSVDACQRLWFDITGMLFYPNNRQIRRSIWWVDLAT 132  
 Qy 148 SKLVQVEIAPHNIAVNATTGNG--ELVSLAVQAIIDRTNTMVIYADEKGLIMYQNSDDSF 206  
 Db 133 DQVLKRFDVPEI---AETGRGLASITVDVKAGCGDAYAVIDLVYRRLVYVYHLNRDRI 189  
 Qy 207 HRLTNTFDYDPRYKLTIVAGESFTVKGICGIALSPV-----TNNLYSPSSHGLYVVD 262  
 Db 190 WSPFHNFNFDPLSGDLSIGGQTFRWDGIFSTLGAOKLDGSRDAYFHPMASTNEFVVS 249  
 Qy 263 TEQFRNQYBENNVOYEGSODILNTQSGKVVSK-----NGVLFGLGVNSGIACVN 314  
 Db 250 N---RVLQQESNAARSDHGDNDFRVLGRSGPSTQSTMHAYDPTGVIFDEIQRNGVGCKW 306  
 Qy 315 EHQVLORESFDVVAQNEETLQMIYSMKIMENLPQSGRINDPEGNEYMLALSNNRMQKIINN 374







; ORGANISM: DROSOPHILA		
US-11-097-143-42198		
Query Match 10.1%; Score 281; DB 6; Length 530;		
Best Local Similarity 22.4%; Pred. No. 3.2e-14;		
Matches 90; Conservative 88; Mismatches 160; Indels 64; Gaps 14;		
Qy	13 LAHSMKVIYEWKHIIDFDGSDERRDAAIKSGEDHTKYNYPFD-----VDRWRDKTFV-TIE 67	
Db	21 LTPGLQVAKQMKLLRYNF-----EPQAPVSPDPNPNQNVLLITGLAVTD-----DRIFVATPK 73	
Qy	68 RNVGVSSLLVNTKKGKGPPLLRPPDWSFAKYE-----DCSGIV--SAFKIADVDFDL 121	
Db	74 LFGVPSTVSWSKAQGDSPFLNAPFDWTFSTNGRSDFNCSDLILTSVYRLVDSNRI 133	
Qy	122 WVLDSGL-----VNNQPMCSPKLTFDLKTSKLVKQVEIP-----HNIAVNATTGM 168	
Db	134 WLDAGISRSLEDEYITCPEKILVLDLATORVVRRIIDFPEVLRGESLFTNMVIDETAK 193	
Qy	169 GELVSLAVQAIDRTNTWYIYIADEKGEGLIMYQNSDDSFHRLTSNTFDYDPRYTKLTVAGE 228	
Db	194 G-----CDDVFYITDTVEGIIIVYDSGKQVTVRSHPMYPDPDFAQSEIHEH 242	
Qy	229 SFTVNGICGIALSPVTNNLYSPSSHGLYYVDTEQFRN---PQYBENNVOYEGSQDIL 285	
Db	243 RFVLMDSGVGLTDFERTGVVYFQPLATDRVFSVHKVNLVLRAGPLPDGKMLDVKLVGKK--- 299	
Qy	286 NTQSGFGKVS-KGVFLGLVGNSGIACVNEHOVLQRESFDVVAQNEETLQMVSMKIME 344	
Db	300 SSQIGLAVSPFSSLLFSPLETAIASWPTTNQ-----SVLAFDRDQLQFVADITTK 355	
Qy	345 NLPQSGRINDPEGNEYMALSNRMQKIINNDFNDFNVNFRIL 386	
Db	356 SEP-----GVYIATASKFHRPFLKNLNFENFRIV 386	
RESULT 13		
US-10-732-923-9918		
; Sequence 9918, Application US/10732923		
; Publication No. US20050108791A1		
; GENERAL INFORMATION:		
; APPLICANT: Edgerton, Michael D		
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES		
; FILE REFERENCE: 38-15(52796)C		
; CURRENT APPLICATION NUMBER: US/10/732,923		
; CURRENT FILING DATE: 2003-12-10		
; PRIOR APPLICATION NUMBER: 10/310,154		
; PRIOR FILING DATE: 2002-12-04		
; NUMBER OF SEQ ID NOS: 24149		
; SEQ ID NO 9918		
; LENGTH: 638		
; TYPE: PRT		
; ORGANISM: Dictyostelium discoideum		
US-10-732-923-9918		
Query Match 10.0%; Score 279.5; DB 5; Length 638;		
Best Local Similarity 24.8%; Pred. No. 5.4e-14;		
Matches 110; Conservative 62; Mismatches 163; Indels 109; Gaps 21;		
Qy	124 LDSGLVNNQP-----MCSPKLLTFDLKTSKLVKQ---VEIPHNIAVNATTGMGELV 172	
Db	233 MDLGSNNLSDDLKALTRQCK-KLKFISLSKCKLTDHGVLELHDCP-----QLM 283	
Qy	173 SLAVQAIDR-TNTWYIYIADEKGEGLIMYQNSDDSFHRLTSNTFDYDPRYTKLTVAGES 229	
Db	284 DLNLASCKVTRTSVQHVQLNHLNLTNLNLS--SPKNLHPTTFPKNPRLNLTTLTIDLS 341	
Qy	230 FT-VKN-----GICGIALSPVTNNLYSPSSH--GLYYVDTEQFRNPOYE 272	
Db	342 FTDVKDEDIFQLTEYAANLKSRLVACVEVTDSEMLM-IATHCKKLCFVDLS--RDQQQS 398	
Qy	273 ENNVYEGSQDILNTQSGFKVSKNGVFLGLVGNSGIACVNEHOVLQRESFDVVAQNEE 332	

; ORGANISM: DROSOPHILA		
US-11-097-143-42198		
Query Match 10.1%; Score 281; DB 6; Length 530;		
Best Local Similarity 22.4%; Pred. No. 3.2e-14;		
Matches 90; Conservative 88; Mismatches 160; Indels 64; Gaps 14;		
Qy	13 LAHSMKVIYEWKHIIDFDGSDERRDAAIKSGEDHTKYNYPFD-----VDRWRDKTFV-TIE 67	
Db	21 LTPGLQVAKQMKLLRYNF-----EPQAPVSPDPNPNQNVLLITGLAVTD-----DRIFVATPK 73	
Qy	68 RNVGVSSLLVNTKKGKGPPLLRPPDWSFAKYE-----DCSGIV--SAFKIADVDFDL 121	
Db	74 LFGVPSTVSWSKAQGDSPFLNAPFDWTFSTNGRSDFNCSDLILTSVYRLVDSNRI 133	
Qy	122 WVLDSGL-----VNNQPMCSPKLTFDLKTSKLVKQVEIP-----HNIAVNATTGM 168	
Db	134 WLDAGISRSLEDEYITCPEKILVLDLATORVVRRIIDFPEVLRGESLFTNMVIDETAK 193	
Qy	169 GELVSLAVQAIDRTNTWYIYIADEKGEGLIMYQNSDDSFHRLTSNTFDYDPRYTKLTVAGE 228	
Db	194 G-----CDDVFYITDTVEGIIIVYDSGKQVTVRSHPMYPDPDFAQSEIHEH 242	
Qy	229 SFTVNGICGIALSPVTNNLYSPSSHGLYYVDTEQFRN---PQYBENNVOYEGSQDIL 285	
Db	243 RFVLMDSGVGLTDFERTGVVYFQPLATDRVFSVHKVNLVLRAGPLPDGKMLDVKLVGKK--- 299	
Qy	286 NTQSGFGKVS-KGVFLGLVGNSGIACVNEHOVLQRESFDVVAQNEETLQMVSMKIME 344	
Db	300 SSQIGLAVSPFSSLLFSPLETAIASWPTTNQ-----SVLAFDRDQLQFVADITTK 355	
Qy	345 NLPQSGRINDPEGNEYMALSNRMQKIINNDFNDFNVNFRIL 386	
Db	356 SEP-----GVYIATASKFHRPFLKNLNFENFRIV 386	
RESULT 14		
US-10-135-322-17		
; Sequence 17, Application US/10135322		
; Publication No. US20020173017A1		
; GENERAL INFORMATION:		
; APPLICANT: BENFEY, PN		
; APPLICANT: HELARIUTTA, Y		
; APPLICANT: MAHONEN, AP		
; APPLICANT: BONKE, AMM		
; APPLICANT: KAUPPINEN, L		
; APPLICANT: RIIKONEN, M		
; TITLE OF INVENTION: WOODEN LEG GENE, PROMOTER AND USES THEREOF		
; FILE REFERENCE: 5914-086-999		
; CURRENT APPLICATION NUMBER: US/10/135,322		
; CURRENT FILING DATE: 2002-04-30		
; PRIOR APPLICATION NUMBER: 60/253,739		
; PRIOR FILING DATE: 2000-11-29		
; NUMBER OF SEQ ID NOS: 43		
; SOFTWARE: PatentIn version 3.0		
; SEQ ID NO 17		
; LENGTH: 2150		
; TYPE: PRT		
; ORGANISM: Arabidopsis thaliana		
US-10-135-322-17		
Query Match 9.5%; Score 267; DB 4; Length 2150;		
Best Local Similarity 20.2%; Pred. No. 3.2e-12;		
Matches 126; Conservative 64; Mismatches 207; Indels 226; Gaps 21;		
Qy	4 NHQRKSANNLAHSMKVIYEWKHIIDFDGSDERR-----DAAIKSGEFDHTKYNYPFD- 54	
Db	235 NNGGNNNNITDPTKSKRHSYETNIGSHQRRKSIQSLIANSIHS--FSKLKNKPLSS 292	
Qy	55 -----VDRWRDKTFVTIER-----NNGVPSLL 76	
Db	293 STPSTVNTCGAVNNNNNNNNNNSTGSLGAIPMDSRFDGNINTITEESTGGNNSPRS- 351	
Qy	77 NVVTNKGKGGPLLRPPDWSFAKVEDCSGIVSAFKIADVDFDLWLDSGLVNN--NQP 134	
Db	352 NCGSNCGSGNGGIPLSP-----RNLSSLNSGVNVSPIRIHL-----NNLNNSNLP 397	
Qy	135 MCSPKLLTFDLKTSKLVKQVEIPHNIAVNATTGMGELVSLAVQAIDRTNTWYIYIADEKGE 194	
Db	398 PLSPRHNFHINVLNLLNN-----NNNINPNNNP----- 427	
Qy	195 GLIMYQNSDDSFHRLT--SNTFDYDPRYTKLTVAGESFTVKGICGIALSPVTNNLYSP 252	
Db	428 -----NNSNNNNVSPNNNNHNISPR-----GSNISPRNNGGSTT 464	
Qy	253 LSSHGLYVDTQEPFRNPOYENNVQYEGSQDIL-----NTQSGFGKVSNGVFLGLVGN 308	
Db	465 ISP-----RNISNNNNIINNINNNTTPRNSPRELVNPNTPNSPRLLATSUNS 513	
Qy	309 GIACVNEHOVLQRESFDVVAQNEBETLQMVSMKIMENLPQSGRINDPEGNEYMALSNRM 368	















[illegible]





Page Blank (up)

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2006, 18:53:48 ; Search time 175.483 Seconds  
(without alignments)  
1034.083 Million cell updates/sec

Title: US-10-525-567-4  
Perfect score: 2178  
Sequence: 1 NLRGSLNKLSPILHEWKF.....NTRCENPDNDRPFKISHL 413

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq 21.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*  
9: Geneseqp2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2177	99.9	413	8 ADL27607	Adl27607 Honey bee
2	2177	99.9	414	8 ADL27617	Adl27617 Honey bee
3	2177	99.9	432	4 AAG78710	Aag78710 Oxygen pr
4	2177	99.9	432	4 AAG78815	Aag78815 Bee prote
5	2177	99.9	432	8 ADL27619	Adl27619 Honey bee
6	1415.5	65.0	524	8 ADL27606	Adl27606 Honey bee
7	1415.5	65.0	544	8 ADL27616	Adl27616 Honey bee
8	1406.5	64.6	544	8 ADL27618	Adl27618 Honey bee
9	452.5	20.8	541	4 ABB60127	Abb60127 Drosophil
10	448	20.6	438	4 ABB60071	Abb60071 Drosophil
11	435.5	20.0	453	4 ABB71124	Abb71124 Drosophil
12	422	19.4	426	4 ABB66147	Abb66147 Drosophil
13	415	19.1	412	4 ABB64733	Abb64733 Drosophil
14	411.5	18.9	415	4 ABB64727	Abb64727 Drosophil
15	366	16.8	429	4 ABB67082	Abb67082 Drosophil
16	358	16.4	530	4 ABB71802	Abb71802 Drosophil
17	357	16.4	379	4 ABB71640	Abb71640 Drosophil
18	343.5	15.8	409	4 ABB66148	Abb66148 Drosophil
19	331.5	15.2	316	4 ABB58294	Abb58294 Drosophil
20	276.5	12.7	397	8 ADO43422	Ado43422 Lutzomyia
21	260	11.9	388	8 ADM16728	Adm16728 PERL-P6-H
22	258	11.8	393	4 ABB57969	Abb57969 Drosophil
23	256	11.8	388	8 ADM16682	Adm16682 PRL-P4-C1
24	240.5	11.0	393	8 ADM16684	Adm16684 PRL-P4-D6

25	233	10.7	382	4 ABB68627	Abb68627 Drosophil
26	229.5	10.5	412	8 ADO43416	Ado43416 Lutzomyia
27	229	10.5	394	6 ABR91192	Abri91192 P. papata
28	214	9.8	393	8 ADM16726	Adm16726 PERL-P7-G
29	209.5	9.6	400	6 ABR91193	Abri91193 P. papata
30	184	8.4	399	8 ADO43456	Ado43456 Lutzomyia
31	156	7.2	342	4 ABB63308	Abb63308 Drosophil
32	138	6.3	25	8 ADL27605	Adl27605 Honey bee
33	123	5.6	1100	2 AAR15783	Aari5783 B.thuring
34	122.5	5.6	575	8 ADS08188	Ads08188 Staphyloc
35	116	5.3	946	5 ABR74063	Abp74063 Candida a
36	114	5.2	414	5 ABB54432	Abb54432 Lactococc
37	113.5	5.2	1530	5 ABB47329	Abb47329 Listeria
38	111	5.1	4643	4 ABB71609	Abb71609 Drosophil
39	110	5.1	391	6 ABU43455	Abu43455 Protein e
40	110	5.1	742	4 ABB67313	Abb67313 Drosophil
41	110	5.1	742	8 ADQ89700	Adq89700 Antagonis
42	110	5.1	1088	4 ABB67108	Abb67108 Drosophil
43	110	5.1	1088	8 ADQ89698	Adq89698 Antagonis
44	109	5.0	4688	6 ABU48941	Abu48941 Protein e
45	108.5	5.0	1427	9 ABM91413	Abm91413 M. xanthu

ALIGNMENTS

RESULT 1  
ADL27607  
ID ADL27607 standard; protein; 413 AA.  
XX ADL27607;  
AC  
XX  
DT 03-JUN-2004 (first entry)  
XX  
DE Honey bee RJP55 mature protein, SEQ ID 4.  
XX  
KW Antiallergic; Antiasthmatic; Antiinflammatory; Dermatological; honey bee;  
KW royal jelly protein; allergy; pollinosis; atopic dermatitis;  
KW contact hypersensitivity; bronchial asthma; allergic rhinitis; RJP55.  
XX  
OS Apis mellifera.  
XX  
PN WO2004019971-A1.  
XX  
PD 11-MAR-2004.  
XX  
PF 26-AUG-2003; 2003WO-JP010795.  
XX  
PR 29-AUG-2002; 2002JP-00252087.  
PR 30-JAN-2003; 2003JP-00022776.  
XX (HAYA/) HAYASHIBARA K.  
XX Okamoto I, Arai N, Kohno K, Kurimoto M, Sano O;  
WPI: 2004-248191/23.  
N-PSDB; ADL27609.  
XX  
PT Antiallergic agents for relieving symptoms accompanying e.g. pollinosis,  
PT atopic dermatitis, contact hypersensitivity, bronchial asthma and  
PT allergic rhinitis, comprise protein originated from royal jelly.  
XX  
PS Claim 2; SEQ ID NO 4; 78pp; Japanese.  
XX  
SQ The present invention relates to novel antiallergic agents, which  
SQ comprise as an active ingredient, peptides ADL27604 and ADL27605, which  
SQ are N-terminal peptides of honey bee royal jelly proteins (RJP, ADL27606  
SQ and ADL27607). The agents can be used to relieve symptoms accompanying an  
SQ allergic disease e.g. pollinosis, atopic dermatitis, contact  
SQ hypersensitivity, bronchial asthma and allergic rhinitis, which are  
SQ applicable in foods, cosmetics and drugs.  
SQ Sequence 413 AA;

Query Match 99.9%; Score 2177; DB 8; Length 413;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-193;  
 Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NILRGESLNKSLPILHEWKFFDYDFGSDERRQDAILSGEYDYKNNYPSDIDQWHDKIFVT 60  
 DB 1 NILRGESLNKSLPILHEWKFFDYDFGSDERRQDAILSGEYDYKNNYPSDIDQWHDKIFVT 60

QY 61 MLRYNGVPSSLNVISKVGGGGLLPQYPDWSPFAKYDDCSGIVSASKLAIDKCDRLWVLD 120  
 DB 61 MLRYNGVPSSLNVISKVGGGGLLPQYPDWSPFAKYDDCSGIVSASKLAIDKCDRLWVLD 120

QY 121 SGLVNNTPQMCSPKLLTFDLTTSQLLKQVEIPHDVAVNATTGKRLSSSLAVQSLDCNTNS 180  
 DB 121 SGLVNNTPQMCSPKLLTFDLTTSQLLKQVEIPHDVAVNATTGKRLSSSLAVQSLDCNTNS 180

QY 181 DTWVYIADEKGEGLIVYHNSDDSFHRLTSNTFDYDPKFTKMTIDGESYTAQDGISGMALS 240  
 DB 181 DTWVYIADEKGEGLIVYHNSDDSFHRLTSNTFDYDPKFTKMTIDGESYTAQDGISGMALS 240

QY 241 PMTNLYSPVASTSLYYNTEQRTSDYQNDIHVEGVQNLDTQSSAKVSKSGVLFF 300  
 DB 241 PMTNLYSPVASTSLYYNTEQRTSDYQNDIHVEGVQNLDTQSSAKVSKSGVLFF 300

QY 301 GLVGSALGWCNEHRTLERHNIRTVASDETLQMIASMKIKEAXPHVPIFDRYINREYIL 360  
 DB 301 GLVGSALGWCNEHRTLERHNIRTVASDETLQMIASMKIKEAXPHVPIFDRYINREYIL 360

QY 361 VLSNKMQKVVNNDFNFDVFRIMNANVNELILNTRCENPDNRTPFKISIH 413  
 DB 361 VLSNKMQKVVNNDFNFDVFRIMNANVNELILNTRCENPDNRTPFKISIH 413

RESULT 2  
 ADL27617  
 ID ADL27617 standard; protein; 414 AA.  
 XX  
 AC ADL27617;  
 XX  
 DT 03-JUN-2004 (first entry)  
 XX  
 DE Honey bee RJP55.  
 XX  
 KW Antiallergic; Antiasthmatic; Antiinflammatory; Dermatological; honey bee;  
 KW royal jelly protein; allergy; pollinosis; atopic dermatitis;  
 KW contact hypersensitivity; bronchial asthma; allergic rhinitis; RJP55.  
 XX  
 OS Apis mellifera.  
 XX  
 PN W02004019971-A1.  
 XX  
 PD 11-MAR-2004.  
 XX  
 PF 26-AUG-2003; 2003WO-JP010795.  
 XX  
 PR 29-AUG-2002; 2002JP-00252087.  
 PR 30-JAN-2003; 2003JP-00022776.  
 XX  
 PA (HAYA/) HAYASHIBARA K.  
 XX  
 PI Okamoto I, Arai N, Kohno K, Kurimoto M, Sano O;  
 XX  
 DR WPI; 2004-248191/23.  
 DR N-PSDB; ADL27609.  
 XX  
 PF Antiallergic agents for relieving symptoms accompanying e.g. pollinosis,  
 PF atopic dermatitis, contact hypersensitivity, bronchial asthma and  
 PF allergic rhinitis, comprise protein originated from royal jelly.  
 XX  
 PS Disclosure; Page 62-65; 78pp; Japanese.  
 XX  
 CC The present invention relates to novel antiallergic agents, which

CC comprise as an active ingredient, peptides ADL27604 and ADL27605, which  
 CC are N-terminal peptides of honey bee royal jelly proteins (RJP, ADL27606  
 CC and ADL27607). The agents can be used to relieve symptoms accompanying an  
 CC allergic disease e.g. pollinosis, atopic dermatitis, contact  
 CC hypersensitivity, bronchial asthma and allergic rhinitis, which are  
 CC applicable in foods, cosmetics and drugs. The present sequence is a honey  
 CC bee RJP.  
 XX  
 SQ Sequence 414 AA;  
 Query Match 99.9%; Score 2177; DB 8; Length 414;  
 Best Local Similarity 99.8%; Pred. No. 2.8e-193;  
 Matches 412; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NILRGESLNKSLPILHEWKFFDYDFGSDERRQDAILSGEYDYKNNYPSDIDQWHDKIFVT 60  
 DB 2 NILRGESLNKSLPILHEWKFFDYDFGSDERRQDAILSGEYDYKNNYPSDIDQWHDKIFVT 61

QY 61 MLRYNGVPSSLNVISKVGGGGLLPQYPDWSPFAKYDDCSGIVSASKLAIDKCDRLWVLD 120  
 DB 62 MLRYNGVPSSLNVISKVGGGGLLPQYPDWSPFAKYDDCSGIVSASKLAIDKCDRLWVLD 121

QY 121 SGLVNNTPQMCSPKLLTFDLTTSQLLKQVEIPHDVAVNATTGKRLSSSLAVQSLDCNTNS 180  
 DB 122 SGLVNNTPQMCSPKLLTFDLTTSQLLKQVEIPHDVAVNATTGKRLSSSLAVQSLDCNTNS 181

QY 181 DTWVYIADEKGEGLIVYHNSDDSFHRLTSNTFDYDPKFTKMTIDGESYTAQDGISGMALS 240  
 DB 182 DTWVYIADEKGEGLIVYHNSDDSFHRLTSNTFDYDPKFTKMTIDGESYTAQDGISGMALS 241

QY 241 PMTNLYSPVASTSLYYNTEQRTSDYQNDIHVEGVQNLDTQSSAKVSKSGVLFF 300  
 DB 242 PMTNLYSPVASTSLYYNTEQRTSDYQNDIHVEGVQNLDTQSSAKVSKSGVLFF 301

QY 301 GLVGSALGWCNEHRTLERHNIRTVASDETLQMIASMKIKEAXPHVPIFDRYINREYIL 360  
 DB 302 GLVGSALGWCNEHRTLERHNIRTVASDETLQMIASMKIKEALPHVPIFDRYINREYIL 361

QY 361 VLSNKMQKVVNNDFNFDVFRIMNANVNELILNTRCENPDNRTPFKISIH 413  
 DB 362 VLSNKMQKVVNNDFNFDVFRIMNANVNELILNTRCENPDNRTPFKISIH 414

RESULT 3  
 AAG78710  
 ID AAG78710 standard; protein; 432 AA.  
 XX  
 AC AAG78710;  
 XX  
 DT 11-DEC-2001 (first entry)  
 XX  
 DE Oxygen promoter related protein.  
 XX  
 KW Oxygen promoter; bee; royal jelly; cardiant; vulnerary;  
 KW cardiopulmonary function.  
 XX  
 OS Apis mellifera.  
 XX  
 PN JP2001172195-A.  
 XX  
 PD 26-JUN-2001.  
 XX  
 PF 15-DEC-1999; 99JP-00356005.  
 XX  
 PR 15-DEC-1999; 99JP-00356005.  
 XX  
 PA (POKK ) POLA CHEM IND INC.  
 XX  
 DR WPI; 2001-592558/67.  
 DR N-PSDB; AAI64996.  
 XX  
 PT Agent for promoting oxygen in cells to improve cardiopulmonary function,  
 PT comprises protein in royal jelly which comprises oxygen enhancing effect.

XX PS Disclosure; Page 9-10; 10pp; Japanese.  
 XX CC The present invention relates to an agent for promoting oxygen,  
 CC comprising a protein in royal jelly which has an oxygen enhancing effect.  
 CC This can be used to promote oxygen in cells, causing an improvement in  
 CC pulmonary function. The present sequence is a protein which can be used  
 CC in the agent of the invention  
 XX SQ Sequence 432 AA;  
 Query Match 99.9%; Score 2177; DB 4; Length 432;  
 Best Local Similarity 100.0%; Pred. No. 3e-193;  
 Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 NILRGESLNKSLPILHEWKPFDDYFGSDERRQDAILSGEYDYKKNYPSDIDQWHDKIFVT 60  
 Db 20 NILRGESLNKSLPILHEWKPFDDYFGSDERRQDAILSGEYDYKKNYPSDIDQWHDKIFVT 79  
 Qy 61 MLRYNGVPSSLNVISKVGDPGLLPYPDWSFAKYDDCSGIVSASKLAIDKCDRLWVLD 120  
 Db 80 MLRYNGVPSSLNVISKVGDPGLLPYPDWSFAKYDDCSGIVSASKLAIDKCDRLWVLD 139  
 Qy 121 SGLVNNTQPMCSPKLLTFDLTTSQLLKQVEIPHDAVNATTKGRLSSLAQSILDCNTNS 180  
 Db 140 SGLVNNTQPMCSPKLLTFDLTTSQLLKQVEIPHDAVNATTKGRLSSLAQSILDCNTNS 199  
 Qy 181 DTWYIADKEGELIVYHNSDDSFRLTNTFDYDPKFTKMTIDGESYTAQDGISGMALS 240  
 Db 200 DTWYIADKEGELIVYHNSDDSFRLTNTFDYDPKFTKMTIDGESYTAQDGISGMALS 259  
 Qy 241 PMTNLYSPVASTSLYYNTEQFRTSDYQNDIHVEGVQNLDTQSSAKVSKGVLPFF 300  
 Db 260 PMTNLYSPVASTSLYYNTEQFRTSDYQNDIHVEGVQNLDTQSSAKVSKGVLPFF 319  
 Qy 301 GLVGSALGCWNEHRTLERHNIRTVQASDETLQMIASMKIKEAXPHVPIDRYINREYIL 360  
 Db 320 GLVGSALGCWNEHRTLERHNIRTVQASDETLQMIASMKIKEAXPHVPIDRYINREYIL 379  
 Qy 361 VLSNKMQRWVNDNFDDVFNFRIMNANVNEILNTRCENPDNDRTPPKISHL 413  
 Db 380 VLSNKMQRWVNDNFDDVFNFRIMNANVNEILNTRCENPDNDRTPPKISHL 432  
 RESULT 4  
 AAG78815  
 ID AAG78815 standard; protein; 432 AA.  
 XX AC AAG78815;  
 XX DT 27-NOV-2001 (first entry)  
 XX DE Bee protein.  
 XX KW Bee; lactic acid accumulation inhibitor; royal jelly; muscular fatigue.  
 XX OS Apis mellifera.  
 XX FH Key Location/Qualifiers  
 XX FT Peptide 1..19  
 XX FT Protein /label= Signal\_peptide  
 XX FT 20..432 /label= Mature\_bee\_protein  
 XX FT Misc-difference 363  
 XX FT /label= unknown  
 XX FT /note= "Encoded by CTV, given as Xaa in specification"  
 XX JP2001172190-A.  
 XX PD 26-JUN-2001.  
 XX PF 15-DEC-1999; 99JP-00356006.  
 XX

PR 15-DEC-1999; 99JP-00356006.  
 XX PA (POKK ) POLA CHEM IND INC.  
 XX WIPI; 2001-592556/67.  
 DR N-PSDB; AAI65041.  
 XX Lactic acid accumulation inhibitor comprises protein contained in royal  
 PT jelly, useful for restoration, improvement and release of muscular  
 PT fatigue.  
 XX Disclosure; Page 9-10; 10pp; Japanese.  
 XX The present invention relates to a lactic acid accumulation inhibitor.  
 CC The inhibitor comprises a protein contained in royal jelly having lactic  
 CC acid accumulation inhibiting activity as the active component. The lactic  
 CC acid accumulation inhibitor is used for restoration, improvement and  
 CC release of muscular fatigue. The present sequence is a bee protein, which  
 CC was used in the present invention  
 XX SQ Sequence 432 AA;  
 Query Match 99.9%; Score 2177; DB 4; Length 432;  
 Best Local Similarity 100.0%; Pred. No. 3e-193;  
 Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 NILRGESLNKSLPILHEWKPFDDYFGSDERRQDAILSGEYDYKKNYPSDIDQWHDKIFVT 60  
 Db 20 NILRGESLNKSLPILHEWKPFDDYFGSDERRQDAILSGEYDYKKNYPSDIDQWHDKIFVT 79  
 Qy 61 MLRYNGVPSSLNVISKVGDPGLLPYPDWSFAKYDDCSGIVSASKLAIDKCDRLWVLD 120  
 Db 80 MLRYNGVPSSLNVISKVGDPGLLPYPDWSFAKYDDCSGIVSASKLAIDKCDRLWVLD 139  
 Qy 121 SGLVNNTQPMCSPKLLTFDLTTSQLLKQVEIPHDAVNATTKGRLSSLAQSILDCNTNS 180  
 Db 140 SGLVNNTQPMCSPKLLTFDLTTSQLLKQVEIPHDAVNATTKGRLSSLAQSILDCNTNS 199  
 Qy 181 DTWYIADKEGELIVYHNSDDSFRLTNTFDYDPKFTKMTIDGESYTAQDGISGMALS 240  
 Db 200 DTWYIADKEGELIVYHNSDDSFRLTNTFDYDPKFTKMTIDGESYTAQDGISGMALS 259  
 Qy 241 PMTNLYSPVASTSLYYNTEQFRTSDYQNDIHVEGVQNLDTQSSAKVSKGVLPFF 300  
 Db 260 PMTNLYSPVASTSLYYNTEQFRTSDYQNDIHVEGVQNLDTQSSAKVSKGVLPFF 319  
 Qy 301 GLVGSALGCWNEHRTLERHNIRTVQASDETLQMIASMKIKEAXPHVPIDRYINREYIL 360  
 Db 320 GLVGSALGCWNEHRTLERHNIRTVQASDETLQMIASMKIKEAXPHVPIDRYINREYIL 379  
 Qy 361 VLSNKMQRWVNDNFDDVFNFRIMNANVNEILNTRCENPDNDRTPPKISHL 413  
 Db 380 VLSNKMQRWVNDNFDDVFNFRIMNANVNEILNTRCENPDNDRTPPKISHL 432  
 RESULT 5  
 ADL27619  
 ID ADL27619 standard; protein; 432 AA.  
 XX AC ADL27619;  
 XX DT 03-JUN-2004 (first entry)  
 XX DE Honey bee MRJP1.  
 XX KW Antiallergic; Antiasthmatic; Antiinflammatory; Dermatological; honey bee;  
 KW royal jelly protein; allergy; pollinosis; atopic dermatitis;  
 KW contact hypersensitivity; bronchial asthma; allergic rhinitis; MRJP1.  
 XX OS Apis mellifera.  
 XX FH Key Location/Qualifiers  
 XX FT Peptide 1..19

```
FT Protein /label= Signal_peptide
FT 20. .432
FT /label= Mature_protein
PN WO2004019971-A1.
XX 11-MAR-2004.
XX
XX 26-AUG-2003; 2003WO-JP010795.
XX
XX 29-AUG-2002; 2002JP-00252087.
XX
XX 30-JAN-2003; 2003JP-00022776.
XX
XX (HAYA/) HAYASHIBARA K.
XX
XX Okamoto I, Arai N, Kohno K, Kurimoto M, Sano O;
XX
XX WPI; 2004-248191/23.
XX
XX N-PSDB; ADL27611.
XX
XX Antiallergic agents for relieving symptoms accompanying e.g. pollinosis,
XX atopic dermatitis, contact hypersensitivity, bronchial asthma and
XX allergic rhinitis, comprise protein originated from royal jelly.
XX
XX Disclosure; Page 70-72; 78pp; Japanese.
XX
XX The present invention relates to novel antiallergic agents, which
XX comprise as an active ingredient, peptides ADL27604 and ADL27605, which
XX are N-terminal peptides of honey bee royal jelly proteins (RJP, ADL27606
XX and ADL27607). The agents can be used to relieve symptoms accompanying an
XX allergic disease e.g. pollinosis, atopic dermatitis, contact
XX hypersensitivity, bronchial asthma and allergic rhinitis, which are
XX applicable in foods, cosmetics and drugs. The present sequence is a honey
XX bee major RJP.
XX
XX Sequence 432 AA;
XX
XX Query Match 99.9%; Score 2177; DB 8; Length 432;
XX Best Local Similarity 100.0%; Pred. No. 3e-193;
XX Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
Qy 1 NILRGESLNKSLPTLHEWKPFDDYDFGSDERRQDAILSGEYDYKKNYPSDIDQWHDKIFVT 60
Db 20 NILRGESLNKSLPTLHEWKPFDDYDFGSDERRQDAILSGEYDYKKNYPSDIDQWHDKIFVT 79
XX
Qy 61 MLRNGVPSSLNVISKVGGGGLLPYPDWSFAKYDDCSGIVSASKLAIDKCDRLWVLD 120
Db 80 MLRNGVPSSLNVISKVGGGGLLPYPDWSFAKYDDCSGIVSASKLAIDKCDRLWVLD 139
XX
Qy 121 SGLVNNTPQMCSPKLLTFLDITTSQLLKQVEIPHDAVNATTGKRLSSLAVQSLDCNTNS 180
Db 140 SGLVNNTPQMCSPKLLTFLDITTSQLLKQVEIPHDAVNATTGKRLSSLAVQSLDCNTNS 199
XX
Qy 181 DTWYIADKGEGLIVHNSDSSFRLTSTNTFDYDPKFTKMTIDGESYTAQDGISGNALS 240
Db 200 DTWYIADKGEGLIVHNSDSSFRLTSTNTFDYDPKFTKMTIDGESYTAQDGISGNALS 259
XX
Qy 241 PMTNLLYSPVASTSLYYVNTQFRTSDYQNDIHVEGVQNLDTQSSAKVSKSVGLFF 300
Db 260 PMTNLLYSPVASTSLYYVNTQFRTSDYQNDIHVEGVQNLDTQSSAKVSKSVGLFF 319
XX
Qy 301 GLVGDALGCWNEHRTLERNIRTVASDFTLQMIASMKIKEAXPHVPIFDRIYNREYIL 360
Db 320 GLVGDALGCWNEHRTLERNIRTVASDFTLQMIASMKIKEAXPHVPIFDRIYNREYIL 379
XX
Qy 361 VLSNMQKQVNNDFNFDVFRIMNVANVELILNTRCENPDNDTPEKISHL 413
Db 380 VLSNMQKQVNNDFNFDVFRIMNVANVELILNTRCENPDNDTPEKISHL 432
XX
XX RESULT 6
XX ADL27606
XX ID ADL27606 standard; protein; 524 AA.
```

```
XX AC ADL27606;
XX XX
XX DT 03-JUN-2004 (first entry)
XX
XX DE Honey bee RJP70 mature protein, SEQ ID 3.
XX
XX KW Antiallergic; Antiasthmatic; Antiinflammatory; Dermatological; honey bee;
XX royal jelly protein; allergy; pollinosis; atopic dermatitis;
XX contact hypersensitivity; bronchial asthma; allergic rhinitis; RJP70.
XX
XX OS Apis mellifera.
XX
XX PN WO2004019971-A1.
XX
XX PD 11-MAR-2004.
XX
XX PF 26-AUG-2003; 2003WO-JP010795.
XX
XX PR 29-AUG-2002; 2002JP-00252087.
XX
XX PR 30-JAN-2003; 2003JP-00022776.
XX
XX PA (HAYA/) HAYASHIBARA K.
XX
XX PI Okamoto I, Arai N, Kohno K, Kurimoto M, Sano O;
XX
XX DR WPI; 2004-248191/23.
XX
XX DR N-PSDB; ADL27608.
XX
XX PT Antiallergic agents for relieving symptoms accompanying e.g. pollinosis,
XX atopic dermatitis, contact hypersensitivity, bronchial asthma and
XX allergic rhinitis, comprise protein originated from royal jelly.
XX
XX PS Claim 1; SEQ ID NO 3; 78pp; Japanese.
XX
XX CC The present invention relates to novel antiallergic agents, which
XX comprise as an active ingredient, peptides ADL27604 and ADL27605, which
XX are N-terminal peptides of honey bee royal jelly proteins (RJP, ADL27606
XX and ADL27607). The agents can be used to relieve symptoms accompanying an
XX allergic disease e.g. pollinosis, atopic dermatitis, contact
XX hypersensitivity, bronchial asthma and allergic rhinitis, which are
XX applicable in foods, cosmetics and drugs.
XX
XX SQ Sequence 524 AA;
XX
XX Query Match 65.0%; Score 1415.5; DB 8; Length 524;
XX Best Local Similarity 65.0%; Pred. No. 2.9e-122;
XX Matches 264; Conservative 63; Mismatches 66; Indels 13; Gaps 2;
XX
Qy 7 SLNKSPLILHEWKPFDDYDFGSDERRQDAILSGEYDYKKNYPSDIDQWHDKIFVTMLRYNG 66
Db 12 NLAHSMKVIYEWKHIDFDFGSDERRDAAIKSGEFDHTKQYFPDVRMRDKTFVTIERNG 71
XX
Qy 67 VPSLNVISKVGGGGLLPYPDWSFAKYDDCSGIVSASKLAIDKCDRLWVLD SGLVNN 126
Db 72 VPSLNVNVNKKGGGGLLPYPDWSFAKYDDCSGIVSASKLAIDKCDRLWVLD SGLVNN 131
XX
Qy 127 TQPMCSPKLLTFLDITTSQLLKQVEIPHDAVNATTGKRLSSLAVQSLDCNTNSDTMVYI 186
Db 132 NQPMCSPKLLTFLDITTSQLLKQVEIPHDAVNATTGKRLSSLAVQSLDCNTNSDTMVYI 188
XX
Qy 187 ADKSGGLIVHNSDSSFRLTSTNTFDYDPKFTKMTIDGESYTAQDGISGMALSPMTNLL 246
Db 189 ADKSGGLIMYQNSDSSFRLTSTNTFDYDPKFTKMTIDGESYTAQDGISGMALSPMTNLL 248
XX
Qy 247 YYSFVASTSLYYVNTQFRTSDYQNDIHVEGVQNLDTQSSAKVSKSVGLFFGLVGDS 306
Db 249 YYSFSLSHGLIYYVDTQFRTSDYQNDIHVEGVQNLDTQSSAKVSKSVGLFFGLVGDS 308
XX
Qy 307 ALGCWNEHRTLERNIRTVASDFTLQMIASMKIKEAXPHVPIFDRIYNREYILVLSNKM 366
Db 309 GIACVNEHQVLRFSFVVAQNEETLQMIVSMKIMENLQSGRINDPEGNEYMLALSNNRM 368
```

Qy	367	QKVNDFNFDVNFVFRIMNANVNLILNTRC-----ENPDN 402	
Db	369	QKIINDFNFDVNFVFRILGANVDDLMRNTKCGRYHNQAGNQADN 414	
RESULT 7			
ADL27616			
ID	ADL27616	standard; protein; 544 AA.	
XX	AC	ADL27616;	
XX			
DT	03-JUN-2004	(first entry)	
XX			
DE		Honey bee RJP70.	
XX			
XX		Antiallergic; Antiasthmatic; Antiinflammatory; Dermatological; honey bee;	
KW		royal jelly protein; allergy; pollinosis; atopic dermatitis;	
KW		contact hypersensitivity; bronchial asthma; allergic rhinitis; RJP70.	
XX			
OS		Apis mellifera.	
XX			
Key		Location/Qualifiers	
Peptide		1..20	
Protein		/label= Signal_peptide	
FT		21..544	
FT		/label= Mature_protein	
XX			
PN	WO2004019971-A1.		
XX			
XX		11-MAR-2004.	
XX			
XX		26-AUG-2003; 2003WO-JP010795.	
PF			
PR	29-AUG-2002; 2002JP-00252087.		
PR	30-JAN-2003; 2003JP-00022776.		
XX			
PA	(HAYA/) HAYASHIBARA K.		
XX			
PI	Okamoto I, Arai N, Kohno K, Kurimoto M, Sano O;		
XX			
DR	WPI; 2004-248191/23.		
DR	N-PSDB; ADL27608.		
XX			
PT	Antiallergic agents for relieving symptoms accompanying e.g. pollinosis,		
PT	atopic dermatitis, contact hypersensitivity, bronchial asthma and		
PT	allergic rhinitis, comprise protein originated from royal jelly.		
XX			
PS	Disclosure; Page 59-62; 78pp; Japanese.		
XX			
CC	The present invention relates to novel antiallergic agents, which		
CC	comprise as an active ingredient, peptides ADL27604 and ADL27605, which		
CC	are N-terminal peptides of honey bee royal jelly proteins (RJP, ADL27606		
CC	and ADL27607). The agents can be used to relieve symptoms accompanying an		
CC	allergic disease e.g. pollinosis, atopic dermatitis, contact		
CC	hypersensitivity, bronchial asthma and allergic rhinitis, which are		
CC	applicable in foods, cosmetics and drugs. The present sequence is a honey		
CC	bee RJP.		
XX			
SQ	Sequence 544 AA;		
Query Match 65.0%; Score 1415.5; DB 8; Length 544;			
Best Local Similarity 65.0%; Pred. No. 3.1e-122;			
Matches 264; Conservative 63; Mismatches 66; Indels 13; Gaps 2;			
Qy	7	SLNKLPTLHEWKFPDYDFGSDERRQDAILSGEYDYKKNYPDSIDQWHDKIFVTMLRYNG 66	
Db	32	NLAHMKVYIYWKHIDPFGSDERRDALKSGEPDKYKYPDVDRWDRKTFVIERNG 91	
Qy	67	VPSSLNVISKVGGGPLLQYPDWSFAKYDCSGIVSASKLAIDKCDRLWVLDGLVNN 126	
Db	92	VPSSLNVVTKKGKGPLLRYPDWSFAKYDCSGIVSFAKIVDKFDRWLWLDGLVNN 151	
Qy	127	TQPMCSPKLLTLDLTTSQLLKQVEIPHVAVNATTGKRLSSLAQSLDCVNSDTMYVI 186	

Db	152	NQPMCSPKLLTLDLTTSKLKQVEIPHVAVNATTGKRLSSLAQSLDCVNSDTMYVI 208	
Qy	187	ADEKGEGLIYVYNSDDSPHRLTSNTFDYDPKFTKMTIDGESYTAQDGISGMALSPMTNNL 246	
Db	209	ADEKGEGLIYVYNSDDSPHRLTSNTFDYDPKFTKMTIDGESYTAQDGISGMALSPMTNNL 268	
Qy	247	YSPVASTSLYYVNTQFRTSDYQONDIIHYEGVQNTLDTQSSAKVYVSKGVLPFGLVGDS 306	
Db	269	YSPVASTSLYYVNTQFRTSDYQONDIIHYEGVQNTLDTQSSAKVYVSKGVLPFGLVGDS 328	
Qy	307	ALGCVNEHRTLEHRTVIRTAQSDTQMTASMKIKKAXPHVPIFDRIYINREYILVLSNKM 366	
Db	329	GIACVNEHQVLRQESDFVVAQNEETLQMIYVSMKIMENLPQSGRINDPEGNEYMLALSNNM 388	
Qy	367	QKVNDFNFDVNFVFRIMNANVNLILNTRC-----ENPDN 402	
Db	389	QKIINDFNFDVNFVFRILGANVDDLMRNTKCGRYHNQAGNQADN 434	
RESULT 8			
ADL27618			
ID	ADL27618	standard; protein; 544 AA.	
XX	AC	ADL27618;	
XX			
DT	03-JUN-2004	(first entry)	
XX			
DE		Honey bee MRJP3.	
XX			
KW		Antiallergic; Antiasthmatic; Antiinflammatory; Dermatological; honey bee;	
KW		royal jelly protein; allergy; pollinosis; atopic dermatitis;	
KW		contact hypersensitivity; bronchial asthma; allergic rhinitis; MRJP3.	
XX			
OS		Apis mellifera.	
XX			
Key		Location/Qualifiers	
Peptide		1..16	
Protein		/label= Signal_peptide	
FT		17..544	
FT		/label= Mature_protein	
XX			
PN	WO2004019971-A1.		
XX			
PD	11-MAR-2004.		
XX			
PF	26-AUG-2003; 2003WO-JP010795.		
XX			
PR	29-AUG-2002; 2002JP-00252087.		
PR	30-JAN-2003; 2003JP-00022776.		
XX			
PA	(HAYA/) HAYASHIBARA K.		
XX			
PI	Okamoto I, Arai N, Kohno K, Kurimoto M, Sano O;		
XX			
DR	WPI; 2004-248191/23.		
DR	N-PSDB; ADL27610.		
XX			
PT	Antiallergic agents for relieving symptoms accompanying e.g. pollinosis,		
PT	atopic dermatitis, contact hypersensitivity, bronchial asthma and		
PT	allergic rhinitis, comprise protein originated from royal jelly.		
XX			
PS	Disclosure; Page 65-69; 78pp; Japanese.		
XX			
CC	The present invention relates to novel antiallergic agents, which		
CC	comprise as an active ingredient, peptides ADL27604 and ADL27605, which		
CC	are N-terminal peptides of honey bee royal jelly proteins (RJP, ADL27606		
CC	and ADL27607). The agents can be used to relieve symptoms accompanying an		
CC	allergic disease e.g. pollinosis, atopic dermatitis, contact		
CC	hypersensitivity, bronchial asthma and allergic rhinitis, which are		
CC	applicable in foods, cosmetics and drugs. The present sequence is a honey		
CC	bee major RJP.		
XX			









```

Db 207 GIVVDVAAQSWRI-ENKTYPHDPFGTGTIAGESFQIWDGTVSTLTTHGLGGRMMY 265
Qy 248 YSPVAS-----TSLYYNTQFRTSDYQNDIHYEGVONILD-----TSSAKVVS 293
Db 266 FHSLSSEWQMAIPLDVN-----NGSNRLND-----VSAALDQQLGKRGSCQVAAAMS 316
Qy 294 KSGVLPGLVGDSSALGTCWNEHRTLERNHRTVAQSDETIQMIASMKIKEAXPHVPIFDY 353
Db 317 ESGFLICGLVQPSALLAWNRTGYSHQNLVNLVEDEQRLQFASGLKIVR-----NH 367
Qy 354 INREYILVLSNMOKMVNDNFDDVNFRIIMNANVELILNTRC 397
Db 368 EGKEELWVLSRLQAFAGAGLDYKEINFRIQKCGVQELLGRPC 411

RESULT 14
ABB64727
ID ABB64727 standard; protein; 415 AA.
XX AC ABB64727;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster polypeptide SEQ ID NO 20973.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX OS Drosophila melanogaster.
XX FN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKE ) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX N-PSDB; ABL08830.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions.
XX PS Disclosure; SEQ ID NO 20973; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
XX ABB72072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 415 AA;
XX Query Match 18.9%; Score 411.5; DB 4; Length 415;
XX Best Local Similarity 28.0%; Pred. No. 5.3e-29;
XX Matches 119; Conservative 74; Mismatches 147; Indels 85; Gaps 15;
Qy 5 GBSLNKSLPI-----LHEWKFFDYDFGSDERRQDAILSGEYDKNNYPSDIDQWH----- 54
Db 19 GVGNGRPRVTRVEILTQWGLEFGFPFPAQDRENAQAAGNLVPENGTFIDVQPYWANGQ 78

```

```

Qy 55 DKLFVTMLRY-NGVPSLNVISKVGGGPLLQPYD--WSFAKYDDCSGIVSASKLAID 111
Db 79 IRLFTTIPFVTGIPYTLATVSATQGRNGELLQPYPNYSWMHANGCEDCDRITSAPRAIT 138
Qy 112 KCDRLWLVDSGLVNNTPCMSPKLLFTDTSOLLKQVEIPHDAVNNATTGKRL-----S 167
Db 139 ECNQMWIDSGVIGTTQ-LCPPLLQFALATDRLLHFRFPNDTYIPS-----GSLFTPN 193
Qy 168 SLAVQSLDCNTNSDTMVYIADKEGGLIVYHNSDDSPHRLTSNTDY-DPKFTQMTIDGE 226
Db 194 VLVDPPPPRGTCRTMIYVADTSWR-----AENRFMPDPDYKHTIAGE 238
Qy 227 SYTAQDGISGMALSPMTNNLYSPVASTSLYYVN-----TEQFRTSY 269
Db 239 SFYLMGDM--FALNDKRNLYFHPLASASEYSVPLSALNRQNWANGPEALPEEFLGR 296
Qy 270 QQNDIHYEGVONILDQTSSAKVSKGVLPFGLVGDSSALGTCWNEHRTLERNHRTVAQSD 329
Db 297 RRS-----ECAAASAIIDGRNNVYCVTFNPKLF-----VMNVNPSYNSRNFGLNPAKS 343
Qy 330 ETLOMTASMKIKEAXPHVPIFDYINR---EYILVLSNMOKMVNDNFDDVNFRIIMNA 386
Db 344 DDLQFVSGMKVLR-----NREGQELMMLSNRYOKIAAGTINSKEVNFRIIRR 391
Qy 387 NVNEL 391
Db 392 KLDDV 396

RESULT 15
ABB67082
ID ABB67082 standard; protein; 429 AA.
XX AC ABB67082;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster polypeptide SEQ ID NO 28038.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX OS Drosophila melanogaster.
XX FN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKE ) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX N-PSDB; ABL11185.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions.
XX PS Disclosure; SEQ ID NO 28038; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
XX ABB72072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 415 AA;
XX Query Match 18.9%; Score 411.5; DB 4; Length 415;
XX Best Local Similarity 28.0%; Pred. No. 5.3e-29;
XX Matches 119; Conservative 74; Mismatches 147; Indels 85; Gaps 15;
Qy 5 GBSLNKSLPI-----LHEWKFFDYDFGSDERRQDAILSGEYDKNNYPSDIDQWH----- 54
Db 19 GVGNGRPRVTRVEILTQWGLEFGFPFPAQDRENAQAAGNLVPENGTFIDVQPYWANGQ 78

```

```
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 429 AA;
Query Match 16.8%; Score 366; DB 4; Length 429;
Best Local Similarity 27.6%; Pred. No. 9.5e-25;
Matches 105; Conservative 68; Mismatches 169; Indels 38; Gaps 12;
Qy 37 SGEYDYKNNYPSDIDQHQHDKIFVTM-LRYNGVPSSLNVI--SKKVGSGPILLQYPDPWSP 93
Db ||| : ||| : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
70 SGSEFIQNNYVPGQVTHFRGRLEFVTPRRQPGIPSTLNYIDLAKDQNSQSPHLRAYPNLAV 129
Qy 94 AKYDDC-SGIVSASKLAIDKCDRLWLDGLV---NNTQPMCSFKLLTFLTTSQLKQV 149
Db :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
130 NQYNASEQNLVSVYRTSVDVCGRLWFVDTGMLEFPNNRQQIRHPSIWIWIDLANDRLLRKF 189
Qy 150 EIPHDVAVNATTGKGRLS-SLAVQSLDCNTNSDTMVYIADKEGGLIVYHNSDDSFHRLT 208
Db ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
190 EIPQSI--VEIGRGLASITIDVGARRCN--DAYAVIPDLVNRRLHVYHLRGSDRIWSFE 243
Qy 209 SNTFDYDPKFTKMTIDGESYTAQDGISGMALSPM----TNNLYSPVASTSLYYVNT--- 261
Db :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
244 HSPFNFDPLSDNLNIGGQTFRWDGIFSATLGSYKPDGSRDVFPHPMASTNEFVVSNRVL 303
Qy 262 -EQFRT--SDYQONDHYEGVQNIQDTSQSAKVVSQGVLFGLVGDPSALGCWNEHRTLE 318
Db :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
304 QQEFNAARSDH-GDDFHLGTRGPGSTQSTWHKYDPRGVIFFAEVQKSGVGCWTKSPFS 362
Qy 319 RHNIRTVAQSDETLQMIASMKIEAPHPVIFDRIYNREYILVLSNKKMQMNVNDFNFD 378
Db :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
363 TENHGSVVSNSSEMIYPSDLTIDE-----EGYIWMNSMMPIFVYSKLDVEK 409
Qy 379 VNFRIANVNELILNTRCE 398
Db |||| : |||
410 YNFRIMRQSTLLAKRGTVCE 429
```

Search completed: May 3, 2006, 19:01:18  
Job time : 179.482 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2006, 19:01:48 ; Search time 31.0175 seconds  
(without alignments)  
1281.133 Million cell updates/sec

Title: US-10-525-567-4  
Perfect score: 2178  
Sequence: 1 NIIRGSLNKLSPILHWEK.....NTRCNPNDRTFPKISHL 413

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_80.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	1347.5	61.9	467	2	S39193	royal jelly protei
2	1216.5	55.9	464	2	S39194	royal jelly protei
3	452.5	20.8	541	2	A25696	yellow protein - f
4	435.5	20.0	453	2	J72523	Yellow-B protein -
5	179.5	8.2	388	2	D75353	yellow-related pro
6	114	5.2	340	2	T22741	hypothetical prote
7	114	5.2	414	2	D86764	hypothetical prote
8	113.5	5.2	1530	2	AH1396	peptidoglycan anch
9	112.5	5.2	1252	2	S21178	botulinum neurotox
10	111	5.1	1088	2	B34106	protein kinase [EC
11	110.5	5.1	1251	2	JH0256	botulinum neurotox
12	110	5.1	2178	2	S58005	alpha-toxin - Clos
13	109.5	5.0	1120	2	H71664	transcription-repa
14	109	5.0	4688	2	F82885	hypothetical prote
15	107.5	4.9	1516	2	E71619	RAD2 endonuclease
16	104	4.8	463	2	T33992	hypothetical prote
17	103.5	4.8	934	2	T08418	protein kinase [EC
18	103.5	4.8	1132	2	H82887	hypothetical prote
19	103	4.7	875	2	S70115	ZIP1 protein - yea
20	103	4.7	3562	2	A47171	chondroitin sulfat
21	102	4.7	446	2	S67437	Damage and replica
22	101	4.6	888	2	E82885	hypothetical prote
23	101	4.6	2229	2	T16199	hypothetical prote
24	100.5	4.6	345	2	A97113	probable Fe-S-clus
25	100.5	4.6	997	2	S73556	MG414 homolog C12
26	100.5	4.6	1162	2	A47708	progenitor toxin n
27	100.5	4.6	2314	1	A46151	protein-tyrosine-p
28	100	4.6	721	2	T05815	hypothetical prote
29	99.5	4.6	589	2	B97806	hypothetical prote

30	99.5	4.6	2496	2	A71616	secreted protein P
31	99	4.5	1112	2	T06307	hypothetical prote
32	98.5	4.5	412	2	C83531	aspartate kinase a
33	98.5	4.5	428	2	S15662	farnesyltransferas
34	98.5	4.5	452	2	F86981	sugar-binding peri
35	98.5	4.5	641	2	E82884	DNA primase UU494
36	98.5	4.5	1217	2	F97177	alpha-glucosidase
37	98.5	4.5	2026	1	OYBY	adenylate cyclase
38	98	4.5	595	2	E82934	hypothetical prote
39	97.5	4.5	796	2	S66769	probable membrane
40	97.5	4.5	1047	2	T14897	probable ATPase (E
41	97.5	4.5	1198	2	T34340	dynamin heavy chain
42	97.5	4.5	1608	2	T17201	adenylate cyclase
43	97.5	4.5	1795	2	F97713	190K antigen precu
44	97.5	4.5	1902	2	C97702	cell surface anti
45	97.5	4.5	2140	2	P95074	serine proteinase,

ALIGNMENTS

RESULT 1

S39193  
royal jelly protein RJP57-1 - honeybee  
C;Species: Apis mellifera (honeybee)  
C;Date: 05-Mar-1995 #sequence\_revision 12-May-1995 #text\_change 12-May-1995  
C;Accession: S39193  
R;Klaudiny, J.; Hanes, J.; Kulifajova, J.; Albert, S.; Simuth, J.  
submitted to the EMBL Data Library, September 1993  
A;Description: Molecular cloning of two cDNAs from the head of the nurse honey bee (Apis mellifera) encoding a protein that is identical to the honeybee royal jelly protein RJP57-1.  
A;Reference number: S39193  
A;Accession: S39193  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-467 <KLA>  
A;Cross-references: UNIPARC:UPI0000179A48; EMBL:Z26318

Query Match 61.9% Score 1347.5; DB 2; Length 467;  
Best Local Similarity 63.3%; Pred. No. 5e-93;  
Matches 257; Conservative 62; Mismatches 72; Indels 15; Gaps 4;

Qy	7	SLNKLPIHLEWKFDFDGSERRQDAIISGEYDYKXNYPSPSIDQWHDKIFVTMLRYNG	66
Db	32	NLAHSMKVIYEWKHIDFDGSDEREM-LRLNLRIDHTKNTY-FDVRWRDKTFVTIERNNG	89
Qy	67	VPSSLNVIKXVGCGPILQYPDPWSFAKYDDCSGIYVSAKLAIDKCDRLWVLSGLVNN	126
Db	90	VPSSLNVTNKKGGKGPYDPWSFAKYDDCSGIYVSAFKIAVDKFDRLWVLSGLVNN	149
Qy	127	TOPMCSKLLTDLTTSOLLKQVEIPHDVAVNATTGKRLSSLAVQSLDCNTNSDTWYI	186
Db	150	NQPMCSKLLTDLTTSOLLKQVEIPHDVAVNATTGKRLSSLAVQSLDCNTNSDTWYI	206
Qy	187	ADEKGEGLIYVHNSDDSFHRLTSNTFDYDPKFTQWTDIGESYTAQDGISGHALSPMTNL	246
Db	207	ADEKGEGLIYVHNSDDSFHRLTSNTFDYDPKFTQWTDIGESYTAQDGISGHALSPMTNL	266
Qy	247	YYSVASTSLYYNTEQFRTSDYQNDIHIEGVQNIIDTQSSAKVSKSVLFFGLVGS	306
Db	267	YYSPLLSHGLYYVDETEQFSNPQYEEENVQVSGSDIILNTQSGFKVSKNGVLFGLVGS	326
Qy	307	ALGCWNEHRTLERINIRTVASDETLQIASMKIKKAPHPVPIFDYINRYEIVLSNKM	366
Db	327	GIACNEHQVLRQSFVDVAQNBEITLQIVSMKINENLQSGRINDEPGEYMLALSNNM	386
Qy	367	QKWNNDNFDDVNFIRIMNANVNLIILNTRC-----ENPDN 402	
Db	387	QKIINDFNFVNFIRILGANVDDLNRNTRCGRYHQNAGNQADN 432	
RESULT 2			
S39194			
royal jelly protein RJP57-2 - honeybee			

C;Species: Apis mellifera (honeybee)  
C;Date: 05-Mar-1995 #sequence\_revision 12-May-1995 #text\_change 09-Jul-2004  
A;Accession: S39194  
R;Klaudiny, J.; Hanes, J.; Kulifajova, J.; Albert, S.; Simuth, J.  
submitted to the EMBL Data Library, September 1993  
A;Description: Molecular cloning of two cDNAs from the head of the nurse honey bee (Apis mellifera) encoding a protein that is homologous to the Drosophila protein.  
A;Reference number: S39193  
A;Accession: S39194  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-464 <KLA>  
A;Cross-references: UNIPROT:Q17061; UNIPARC:UPI000012F56A; EMBL:Z26319; NID:9433530; PID:9433530

Query Match 55.9%; Score 1216.5; DB 2; Length 464;  
Best Local Similarity 58.1%; Pred. No. 3.1e-83;  
Matches 233; Conservative 71; Mismatches 84; Indels 13; Gaps 6;

QY 5 GESLNKSLPILHEMKPFDFGSDERRQDAILSGEYDKKNYPSDIDQWHDKIFVTMLRY 64  
DB 29 GKMLTNTLNVHKKYLDYDFDNDERQAAIQSGEYDRTKNYPDLDVQWKNKTFILAVIRY 88  
QY 65 NGVPSSLNVIKSKVGDGGLLPYPDWSFAKYDDCSGIVSASKLAIDKCDRLMWLDSGLV 124  
DB 89 NGVPSSLNVIKSKVGDGGLLPYPDWSFAKYDDCSGIVSASKLAIDKCDRLMWLDSGLV 148  
QY 125 NNTQPMCSKLLTFDLTTSQLLKQVEIPHDVAVNATTKGRLSSLAQSLDCNNTSDTMV 184  
DB 149 NNTQPMCSKLLTFDLTTSQLLKQVEIPHDVAVNATTKGRLSSLAQSLDCNNTSDTMV 202  
QY 185 YIADEKGEGLIVVHNSDDSPHRLTSTNTFDYDPKFTKMTIDGESYTAQD---GISGVALSP 241  
DB 203 YMDVKNKT-LIIVQNAADSSPHRLTSTNTFDYDPKFTKMTIDGESYTAQD---GISGVALSP 259  
QY 242 MTNNLYSPVASTSLYYVNTTEQFRTSDYQNDIHVEGVQNLDTQSSAKVSKSGLVFFG 301  
DB 260 VTHNLYNPSSENLYVNTESLMSKSENQNDVQYERQVDFDSQLTKAVSKNGVLLFG 319  
QY 302 LVGDSALGCVNNEHRTLRHNRITVAQSDTLQMIASMKIKKXAPHPVIFDRIYINREYILV 361  
DB 320 L-ANNTLSKNEHQSLDRQIDVVARNEEDLQMWVSMKIKQNVPPQSGRVNNTQORNEYLLA 378  
QY 362 LSKNMQKQVNVNDFDVFNRIMNANVVELILNTRCENPDN 402  
DB 379 LSDRNQNVLANDLNLEHVNFEQLGANVNDLIRNSRCANFDN 419

RESULT 3  
A25696  
Yellow protein - fruit fly (Drosophila melanogaster)  
C;Species: Drosophila melanogaster  
C;Date: 21-May-1988 #sequence\_revision 21-May-1988 #text\_change 09-Jul-2004  
A;Accession: A25696; A25683  
R;Gevers, P. K.; Spana, C.; Corces, V. G.  
EMBO J. 5, 2657-2662, 1986  
A;Title: On the molecular mechanism of gypsy-induced mutations at the yellow locus of Drosophila.  
A;Reference number: A25696; MUID:87053835; PMID:3096713  
A;Accession: A25696  
A;Molecule type: DNA  
A;Residues: 1-541 <GEY>  
A;Cross-references: UNIPROT:P09957; UNIPARC:UPI000011F616; GB:X04427; NID:g8835; PIDN:CA  
R;Chia, W.; Howes, G.; Martin, M.; Meng, Y. B.; Moses, K.; Taubota, S.  
EMBO J. 5, 3597-3605, 1986  
A;Title: Molecular analysis of the yellow locus of Drosophila.  
A;Reference number: A25683  
A;Accession: A25683  
A;Molecule type: DNA; mRNA  
A;Residues: 1-541 <CHI>  
A;Cross-references: UNIPARC:UPI000011F616; GB:X04703  
C;Genetics:  
A;Gene: FlyBase:Y  
A;Cross-references: FlyBase:FBgn0004034  
A;Introns: 80/1

Query Match 20.8%; Score 452.5; DB 2; Length 541;  
Best Local Similarity 30.3%; Pred. No. 5.4e-26;  
Matches 125; Conservative 90; Mismatches 150; Indels 47; Gaps 17;

QY 16 HEWKPFDFGSDERRQDAILSGEYDKKNYPSDIDQWHDKIFVTMLRY-NGVPSSLNVI 74  
DB 29 YSWSQLDFAPNTRLKQALASGDYIPQNALPGVHEFGNRLFVTVPRWRDGLPATLYI 88  
QY 75 S-KKVGDDGGLLPYPDWSFAKYDDC-SGIVSASKLAIDKCDRLMWLDSGLV---NNTQ 129  
DB 89 NMDRSITSGPELIPYDWSNTAGDCANSITTAIRIKVDECGRLMWLDTGTGIGNTTN 148  
QY 130 MSCP KLLTFDLTTSQLLKQVEIPHDVAVNATTKGRLSSLAQSLDCNNTSDTMVYIAD 188  
DB 149 PCPYANVFDLTDTTRIRYELP-GVDTNPNT---FIANIV---DIGKNCDDAYAFAD 201  
QY 189 EKGEGLIVVHNSDDSPHRLTSTNTFDY-DPKFTKMTIDGESYTAQD---GISGVALSP 245  
DB 202 ELGYGLIAYSWELNKSFRSAHSYFPDPLRGDFNAGINFQWGEIGITGMSLSPRS 261  
QY 246 ---LYTSPVASTSLYYVNT---EQFRTSDYQNDIHVEGVQNLDTQSSAKVSKSGLV 298  
DB 262 YRTLYESPLASHRQPAVSTRILRDETRTDSYHDFVALD--ERGPNSHTTSRVMSDDG 319  
QY 299 FGLVGDGSLGCVNNEHRTLRHNRITVAQSDTLQMIASMKIKKXAPHPVIFDRIYI 354  
DB 320 LFNLDIQNAVGCW--HSSMPYSPQPHGI--VDRDDVGLVFPADVKIDE----- 363  
QY 355 NREYILVNLKQKQVNVNDFDVFNRIMNANVVELILNTRCENPDNDRTP 406  
DB 364 -KNVNLSDRMPVFLSLDLDISDTNFRITAPLATIENTVCDLRNAYGP 414

RESULT 4  
JC7253  
Yellow-B protein - fruit fly (Drosophila melanogaster)  
C;Species: Drosophila melanogaster  
C;Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 09-Jul-2004  
A;Accession: JC7253  
R;Maleszka, R.; Kucharski, R.  
Biochem. Biophys. Res. Commun. 270, 773-776, 2000  
A;Title: Analysis of Drosophila yellow-B cDNA reveals a new family of proteins related to rans.  
A;Reference number: JC7253  
A;Accession: JC7253  
A;Molecule type: mRNA  
A;Residues: 1-453 <WAL>  
A;Cross-references: UNIPROT:Q9VJI5; UNIPARC:UPI0000083EA4; GB:AC005119  
A;Experimental source: head  
C;Comment: This protein, a secreted protein, is involved in neural function, cuticle development.  
A;Gene: yellow-B  
A;Map position: 2 left arm

Query Match 20.0%; Score 435.5; DB 2; Length 453;  
Best Local Similarity 28.4%; Pred. No. 7.7e-25;  
Matches 116; Conservative 85; Mismatches 165; Indels 43; Gaps 12;

QY 9 NKSPLTHLHWKPFDFGSDERRQDAILSGEYDKKNYPSDIDQWHDKIFVTMLRY-NGV 67  
DB 19 NDNLVAYEWREMDFKYANPDQWQSAIERGEKFPANVIFGLEVAGHRLVFLPRWRDGV 78  
QY 68 PGS-----LNVISKKVGDGGLLPYPDWSFAKYDDCS-GIVSASKLAIDKCDRLMWL 120  
DB 79 PASLAVLDLNDTSSK---GPALKPPPSQAHNLQEAPELVSPFRVADRCCGRLWVLD 134  
QY 121 --SGLVNTQPMCSKLLTFDLTTSQLLKQVEIPHDVAVNATTKGRLSSLAQSLDCN 178  
DB 135 RISGVLEQTKIYGAAQLVYDLHNDLLRR---HVLPAQLKQGSLLANLAVEDSDC-- 188  
QY 179 NSDTMVIYIADEKGEGLIVVHNSDDSPHRLTSTNTFDYDPKFTKMTIDGESYTAQD 238

Db 189 -ENTFAAADLGSPGLVVISKDBEESWRVQHFFHPDPMAGNSINGIEFQWDDGLYGLA 247  
 Qy 239 LS-PMTN---NLXYSPVASTSLYYVANTEQFTSDYQQNDIHYEGVQNI-----LDTOSSAK 290  
 Db 248 LSKPLETGATLYFHPCLCSITEFSVDTSLURNKTLATSPMIYREFKVLGSRGPTQAGAE 307  
 Qy 291 VVS-KSGVLFGLVGDSALCGCWEHRTLERHNIRTVAAQSDETLQMTASMKIEXAPHVPI 349  
 Db 308 FLDPDGTGVLFPALPNLNEVACWRTATDFSHSOSQRHMMNNDTLVFPFSDIKVDD----- 360  
 Qy 350 FDRYINREYILVLSNKKMKVNNDDVDVNFRIIMANVANVELIINTRCE 398  
 Db 361 -----QKRLWVLSNQLPVFIYDELYAGSINFRILTASVKEAIENTACE 403

RESULT 5  
 D75353  
 yellow-related protein - Deinococcus radiodurans (strain R1)  
 C:Species: Deinococcus radiodurans  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
 C:Accession: D75353  
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
 M.; Shen, M.; Yamachavan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A:Reference number: A75250; MUID:20036896; PMID:10567466  
 A:Accession: D75353  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-388 <WHI>  
 A:Cross-references: UNIPROT:Q9RTH4; UNIPARC:UPI00000D3ED5; GB:AE000513; NID  
 A:Experimental source: strain R1  
 C:Genetics:  
 A:Gene: DR1790  
 A:Map position: 1

C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T22741  
R;Baynes, C.  
submitted to the EMBL Data Library, November 1996  
A;Reference number: Z19608  
A;Accession: T22741  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-340 <WILL>  
A;Cross-references: UNIPROT:Q9XUE1; UNIPARC:UPI00000790CC; EMBL:Z82272; PIDN:CAB05215.1;  
A;Experimental source: clone F55G11  
C;Genetics:  
A;Gene: CRSP:F55G11.6  
A;Map position: 4  
A;Introns: 21/1; 97/2; 134/1; 200/2; 239/1; 287/3; 308/2

Query Match 5.2%; Score 114; DB 2; Length 340;  
Best Local Similarity 24.3%; Pred.No.0.56; Gaps 18;  
Matches 76; Conservative 33; Mismatches 98; Indels 106; Gaps 18;

Qy 72 NVISKVGCGPGLLPYDWS-----PAKYDDCSGIVSA-----SKLAI-----DKCD 114  
Db |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||  
Qy 26 NVINKPVIDGPVWP-PWTETQPAPQLDKQSQCWIPTIRGYAKLIISGKTTDKDS 84  
Db |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||  
Qy 115 RLWLVDLS--GLVNNTQPMCPKLL---TPDLTTSQLLKQVEPHDVAVNATTCGRUSSL 169  
Db |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||  
Qy 85 RFQTVDASAGNIQTTOENMDPYFFPKASKFLAVS-----NEGSATFAFK 128  
Db |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||  
Qy 170 AVQSLDCNTSDMTVIYIADEKG---EGLIV-----YNSDDSFHRLTSTPDYDP 216  
Db |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||  
Qy 129 VWFEL-----LPYVGDDAGIGPSGLVVNVTSKYVAIDYHSTDG---LLTLMTFPADT 177  
Db |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||  
Qy 217 KFTKMTIDESYTAODGI--SCWALSPMTNNLYSPVASTLSYYVANTEQERTSYQQNDI 274  
Db |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||  
Qy 178 K-----NYSLRSALVFAGEGLSNYYTNLY-----LLYTQKKPWISSDDAITV 221  
Db |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||  
Qy 275 HVEGV---QNILDTQSS-----AKVYSKSGVLFFGLVGDSALGCWNHERTLRRHI 322  
Db |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||  
Qy 222 NFEARASSNDNL-L-IQSSKYLITGIGEWELHPQANSIYNGTVNGTGL-----MS 268  
Db |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||  
Qy 323 RTVAQSDTELQMI 335  
Db |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||  
Qy 269 SILVAVTDLEPMQMI 281  
Db |||:|||:|||:|||:|||:|||:|||:|||:|||:||||||

RESULT 7  
D86764  
hypothetical protein ylcF [imported] - Lactococcus lactis subsp. lactis (strain IL1403)  
C;Species: Lactococcus lactis subsp. lactis  
C;Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 09-Jul-2004  
C;Accession: D86764  
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jallou, O.; Malarme, K.; Weissenbach, J.; Ehrlich  
Genome Res. 11, 731-753, 2001  
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss  
A;Reference number: AB6625; MUID:21235186; PMID:11337471  
A;Accession: D86764  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-414 <STO>  
A;Cross-references: UNIPROT:Q9CGI0; UNIPARC:UPI0000004431; GB:AE005176; PID:g12724076; F  
A;Experimental source: strain IL1403  
C;Genetics:  
A;Gene: ylcF

T22741  
hypothetical protein F55G11.6 - *Caenorhabditis elegans*



Qy 280 -QNILDTQSSAKVYKSGVLFPGLVGDSDL-----GCWNE----- 313  
Db 260 RENILCAGSSR---GGMGALYHGLGVALVSMDDPVVDRSPWLQSDADVQMLFDCIPVSFV 316  
Qy 314 --HRTLEHNRITVAOSDETFLOMIASMKIEAAXPHPIFDYIRNRVILVLSNKMOKV 370  
Db 317 DKLNLLEKTNL-----SAEKIQVITS-----PQPIYTFPIQIKWKLAKTYRMK 364  
Qy 371 NNDNFDP-----DVNFRIMNANVNELNTRCENPDN 402  
Db 365 LTDRQFDYQPYGGRKHGDFVNRNIPLLMLKINEFLYG--CDSIEN 407  
RESULT 8  
AH1396  
peptidoglycan anchored protein (LPXTG motif) [imported] - Listeria monocytogenes (strain  
C;Species: Listeria monocytogenes  
C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C;Accession: AH1396  
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,  
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A;Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma  
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,  
A;Title: Comparative genomics of Listeria species.  
A;Reference number: AB1077; MUID:21537279; PMID:11679669  
A;Accession: AH1396  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1530 <GLA>  
A;Cross-references: UNIPROT:Q8Y479; UNIPARC:UPI0000054CB4; GB:NC\_003210; PIDN:CAD00654.1  
A;Experimental source: strain BGD-e  
C;Genetics:  
A;Gene: lmo2576  
Query Match 5.2%; Score 113.5; DB 2; Length 1530;  
Best Local Similarity 20.8%; Pred. No. 5.6;  
Matches 106; Conservative 79; Mismatches 182; Indels 143; Gaps 32;  
3 LRGSLSNKSPLLHEW-KFFDYDFGSDERRQDAI-----LSGEYDYKKN----- 45  
522 LNGDS-----SWKGFQINFNRDITGQHVITYKTQNPSTHTSGDNEYTNNAIKT 572  
Qy 46 ---YPSDID-QWHDKI FVTMLRYNGVFPSSLNISKVGDGGLQPYDPDMSFAKYDCSG 101  
Db 573 DTAEASDSTKWIDKVIDADGCKNG-----VFNYKTGE-----IEWKLI-FNDSKK 617  
Qy 102 IVSASKLAID---KCDRLWVLDSGLVNTQPMCSPKI-----LTFDLT-----TSQL 145  
Db 618 LI--SKPTIEDSLNSGQTFTQDSIEIHKIDLSATPQVGELIPPENYDVTFTKENGNEQM 675  
Qy 146 LKQVEIP--HDVAVNATT-----GKGRLSLAVQSLDCNTNSDTMWY 185  
Db 676 LITFKKPLIHPEVITYKTKPVGIIKPKYKNAKAVISDGEVLADYEAVIDNANK-----Y 731  
Qy 186 I---ADEKGEGL---IVYHNSDPSHRLTSNTFDYDPKTKMTIDGES---YTAQDGISG- 236  
Db 732 VNKSGEQVDNIDWEIVANQSGS---TVSNATVTDTLTGQKLDTSIIKVKYSQTSVTGK 788  
Qy 237 -----MALSPMTNLYSPVASTSLYVNTQFRTSDYQNDIHYE-GVQNILDTOSSA 289  
Db 789 MLOESNMPISPGEYDLTKTGVDSEENLEYFQV-KFKBNINOSYVIKYQTATLTLSDTETTA 847  
Qy 290 KVKSGSVLPFGLVGDSDLGALGCWNHRLERHIRT-VAQSD-----ETLOWIASMKIKEA 343  
Db 848 QI--GNSVTF---TGDNI-----TKGETEKNIEVKITGDTGCTGTGKILN-KYDKA 897  
Qy 344 XPHVPI-----FDYINREYILVLNKMOKMVN--NDFNPDVNFIMNANVNELILNTRCE 398  
Db 898 DPSIPLEGATFDLVANDEKVDVTQTTDKNGVIEFDLVGYDVTLKEVSAPEGYTLPTASTE 957  
Qy 399 NPD-----NDRTPFKLS--IHL 413

Db 958 NIOVKLEQDEKVVQVWNRKMPKETGEVHL 987  
RESULT 9  
S21178  
botulinum neurotoxin type E precursor - Clostridium botulinum  
C;Species: Clostridium botulinum  
C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 31-Dec-2004  
C;Accession: S21178; JH08107; JH05294; A60027; S18111  
R;Whelan, S.M.; Elmore, M.J.; Bodsworth, N.J.; Atkinson, T.; Minton, N.P.  
Eur. J. Biochem. 204, 657-667, 1992  
A;Title: The complete amino acid sequence of the Clostridium botulinum type-E neurotoxin,  
A;Reference number: S21178; MUID:92174922; PMID:1541280  
A;Accession: S21178  
A;Molecule type: DNA  
A;Residues: 1-1252 <WHE>  
A;Cross-references: UNIPROT:Q00496; UNIPROT:Q45862; UNIPARC:UPI00000010A3; EMBL:X62683; J  
R;Campbell, K.D.; Collins, M.D.; East, A.K.  
J. Clin. Microbiol. 31, 2255-2262, 1993  
A;Title: Gene probes for identification of the botulin neurotoxin gene and specific ide  
A;Reference number: S48103; MUID:94013372; PMID:8408542  
A;Accession: S48107  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 616-982 <CM>  
A;Cross-references: UNIPARC:UPI000008CF0; EMBL:X70815; NID:9407786; PIDN:CAA50146.1; PI  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993  
R;Poulet, S.; Hauser, D.; Quanz, M.; Niemann, H.; Popoff, M.R.  
Biochem. Biophys. Res. Commun. 183, 107-113, 1992  
A;Title: Sequences of the botulin neurotoxin E derived from Clostridium botulinum type  
A;Reference number: JH0256; MUID:92181428; PMID:1543481  
A;Accession: JH0257  
A;Status: nucleic acid sequence not shown  
A;Molecule type: DNA  
A;Residues: 1-176, 'R', 178-197, 'C', 199-339, 'R', 341-772, 'I', 774-962, 'FE', 965-966, 'R', 968-11  
A;Cross-references: UNIPARC:UPI000016EA7F; EMBL:X62089; NID:940393; PIDN:CAA43999.1; PID  
A;Experimental source: strain Beluga  
R;Binz, T.; Kurazono, H.; Wille, M.; Frevert, J.; Wernars, K.; Niemann, H.  
J. Biol. Chem. 265, 9153-9158, 1990  
A;Title: The complete sequence of botulinum neurotoxin type A and comparison with other  
A;Reference number: A35294; MUID:90264400; PMID:2160960  
A;Accession: B35294  
A;Status: not compared with conceptual translation  
A;Molecule type: DNA  
A;Residues: 1-176, 'R', 178-197, 'C', 199-339, 'R', 341-772, 'I', 774-962, 'FE', 965-966, 'R', 968-11  
A;Cross-references: UNIPARC:UPI000017670F  
A;Experimental source: strain Beluga  
R;Gimenez, J.A.; DasGupta, B.R.  
Biochimie 72, 213-217, 1990  
A;Title: Botulinum neurotoxin type E fragmented with endoproteinase Lys-C reveals the sit  
A;Reference number: A60027; MUID:90344918; PMID:2116911  
A;Accession: A60027  
A;Molecule type: protein  
A;Residues: 420-427 <GIM>  
A;Cross-references: UNIPARC:UPI0000176710  
A;Experimental source: strain Beluga  
A;Note: this fragment was generated by proteolysis with Lys-C rather than with trypsin  
C;Comment: The clostridial neurotoxins are highly potent protein toxins that inhibit neu  
C;Comment: The heavy chain mediates the binding of toxin to cell receptors while the lig  
C;Keywords: neurotoxin  
F;1-427/Product: botulinum neurotoxin type E light chain #status predicted <LCH>  
F;423-1252/Product: botulinum neurotoxin type E heavy chain #status predicted <HCH>  
F;412-426/Disulfide bonds: #status predicted  
Query Match 5.2%; Score 112.5; DB 2; Length 1252;  
Best Local Similarity 20.3%; Pred. No. 4.9;  
Matches 96; Conservative 83; Mismatches 173; Indels 121; Gaps 26;  
Qy 1 NILRGSLNKSPLI-----LHEWKFYDFGSDERRQALLSGEYDYKKNYPSDI 50  
Db 414 NIVSVKGIRKSICIEINNGELFFVASENSYNDONINPTKEIDDTVTS-----NNNYENDL 468



```
QY 51 DQHKDFVTLRYNGVPSSLNVIKVGDPGLLOPYPDWSF-----AKY-DDCGSIVSASKLAI-DKCDR 115
DB 469 DQ-----VILFNS-ESAPGLSDEKLN-----LTIONDAYIPKYDS-NGTSDIBQHDV 514
QY 111 DKCDRLVLDGLV-----NN-----TOPMCSPKLLTFDITTSQLLKQVEIP----- 152
DB 515 NELNVFYLDAQKVPGENNVNLTSSIDTALLQPKIYTF--FSSEFINNVNKPVQAALF 572
QY 153 -----HDVANNATGKGRLS-----SLAVQSILDCNTNSDTWVYIADKGBGLIVHNS 200
DB 573 VSMIQQLVDFTEANQKSTVDKIADISIVVPYIGLALN-----IGNEAQKG-----NF 621
QY 201 DDFSHRLTSN--TFDYDPKF---TKWTIDGESYTAQDGIAGSMALSPMTNLL-----Y 247
DB 622 KDALELIGAGLLEFEPELLIPTLVFTIKSFSSDNKNKVKAIKAINNAKERDEKKEV 681
QY 248 YSPVASTSLYVNTQPRTSYQONDHYEGVQNILDTQSSAKVSKSGVLFFGLVGDSA 307
DB 682 YSFIVSNWMTKINT-QFNKRKEQM-----YQALQNV---NAIKTILIESK----- 722
QY 308 LGCWNEHRTLERNHIRTVAQSDFTLQMIASMKIKEAXPHVPIPDYINR---EYILVLSN 364
DB 723 ---YNSY-TLEEKNELTKYDIKOIENELNAKQVSIAMNNI---DRFLTSSISYLMKLN 775
QY 365 ---KMKQVNVNDFPDD--VNPRIMNVN---NELILNTRCENPDNDRTPEKIS 410
DB 776 EVKINKLREYDENVKTYLLNYIIOHGSILGESQQLNSMVTDTLNNISIPFKLS 828

RESULT 10
B34106
protein kinase (EC 2.7.1.37), cGMP-dependent 2, type 1 [similarity] - fruit fly (Drosophila)
N;Contains: protein kinase, cGMP-dependent 2, type 3; protein kinase, cGMP-dependent 2,
C;Species: Drosophila melanogaster
C;Date: 05-May-2000 #sequence revision 05-May-2000 #text_change 09-Jul-2004
C;Accession: B34106; E34106; F34106
J. Kalderson, D.; Rubin, G.W.
R. Biol. Chem. 264, 10738-10748, 1989
A;Title: cGMP-dependent protein kinase genes in Drosophila.
A;Reference number: A34106; MUID:89278147; PMID:2732245
A;Accession: B34106
A;Molecule type: DNA; mRNA
A;Residues: 1-1088 <KAL>
A;Cross-references: UNIPROT:Q03043; UNIPARC:UPI000012DDC5; GB:M27117; GB:M27118; GB:M27119
A;Note: The authors translated the codon AGG for residue 692 as Thr
A;Accession: E34106
A;Status: not compared with conceptual translation
A;Molecule type: DNA; mRNA
A;Residues: 347-507, 'S', 508-691, 'T', 693-1088 <KA2>
A;Cross-references: UNIPARC:UPI000017CE1A
A;Accession: F34106
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA; mRNA
A;Residues: 521-691, 'T', 693-1088 <KA3>
A;Cross-references: UNIPARC:UPI000017CE1A
C;Genetics:
A;Gene: FlyBase:for
A;Cross-references: FlyBase:FBgn0000721
C;Keywords: alternative splicing; ATP: cGMP binding; phosphotransferase; serine/threonine
F;347-1088/Product: protein kinase, cGMP-dependent 2, type 3 #status predicted <TV3>
F;520-637/Domain: cAMP receptor protein cyclic nucleotide-binding domain homology <CAP1>
F;521-1088/Product: protein kinase, cGMP-dependent 2, type c4 #status predicted <CD4>
F;638-762/Domain: cAMP receptor protein cyclic nucleotide-binding domain homology <CAP2>
F;775-1036/Domain: protein kinase homology <KIN>
F;783-791/Region: protein kinase ATP-binding motif
Query Match 5.1%; Score 111; DB 2; Length 1088;
Best Local Similarity 20.6%; Pred. No. 5.2;
Matches 103; Conservative 68; Mismatches 156; Indels 172; Gaps 25;
QY 3 LRGESLNKSLPILHEWKKEFDYDFGSDERRQDAILSGEYDYKNYPSDIDQWHDKIFVTML 62
DB 481 ISAEQSSSSLLHEVSPFKYD--KDERSRELIIKAIILD--NDFMKNLDLTQIREIVDCM 536
```

```
QY 63 RYNGVPSSLNVIKVGDPGLLOPYPDWSF-----AKY-DDCGSIVSASKLAI-DKCDR 115
DB 537 YPKVPAK-NLIKE-GDVGSIYVWEDGRVEVREGKYLSTLSGAKVLGELAILYNCQR 594
QY 116 -----LMVLD-----SGLVNTQPMCSPKLLTFDITTSQLLKQVEIPHD 154
DB 595 TATITAITECNLWAIERQCQFQTIMMRTGLIROAE-----YSDFLKSVPIFKD 641
QY 155 VAVNATTKGRSLSSLAQSLDCNTNSDTWVYIAD-----EKGEGLIVVHNSDDSPH-- 205
DB 642 LA-----EDTLIKISDLVEETHYQRGDHIVROGARGDTFFII 678
QY 206 -----RLTSNTFD-YDPKFTKMTIDGESY-----TAODGISGMALSPM 242
DB 679 SKKRVVITIKQDROEKPIRMGLKGDFFGEGKALQGDLLRTANICESADGVSCLVIRE 738
QY 243 TNNLYSPVASTSLYVNV-----TEOFRSDYQONDHYEGVQNILDTQSSAKVVS 293
DB 739 TFNQLISNLDEIKHRYDDEGAMERRKINEFR-----DI-----NLTD----LRVIA 781
QY 294 KSGVLFFGLV-----GDSALGCWNEHRTLERNHIRTVAQSDFTLQMIASMKIKEAXPH- 346
DB 782 TLGVGGFGRVELVQTNQDSS-----RSPALKQMKK-SQIVETROQQOHIMSEKKEINGEA 833
QY 347 -----VPIPDYINREYILV-----LSNKMQKQVNVNDFNDDVNFRIIMNVNNE----- 390
DB 834 NCOFIKVLFTFKDKKYLMLMESCLGGELWTILRDKGNFDDSTTRFYTACVVEAFDYLH 893
QY 391 -----LILNTR 396
DB 894 SRNIIVRLKAPENILLNER 912

RESULT 11
JH0256
botulinum neurotoxin type E precursor - Clostridium butyricum
C;Species: Clostridium butyricum
C;Date: 30-Jun-1992 #sequence revision 15-May-1998 #text_change 09-Jul-2004
C;Accession: JH0256; S16145
R;Poulet, S.; Hauser, D.; Quanz, M.; Niemann, H.; Popoff, M.R.
Biochem. Biophys. Res. Commun. 183, 107-113, 1992
A;Title: Sequences of the botulinum neurotoxin E derived from Clostridium botulinum type
A;Reference number: JH0256; MUID:92181428; PMID:1543481
A;Accession: JH0256
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-27, 'E', 29-1251 <POU>
A;Cross-references: UNIPROT:P30995; UNIPARC:UPI000017670D; EMBL:X62088; NID:G40379
A;Experimental source: strains ATCC 43181 and ATCC 43755
R;Fujii, N.; Kimura, K.; Yashiki, T.; Indoh, T.; Murakami, T.; Tezuka, K.; Yokosawa, N.
J. Gen. Microbiol. 137, 519-525, 1991
A;Title: Cloning of a DNA fragment encoding the 5'-terminus of the botulinum type E toxin
A;Reference number: S16145; MUID:91237316; PMID:2033376
A;Accession: S16145
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-229, 'M', 231-252 <FUJ>
A;Cross-references: UNIPARC:UPI000016E8AF; EMBL:X53180; NID:G40407; PIDN:CAA37321.1; PID
A;Experimental source: strain BL6340
C;Comment: The clostridial neurotoxins are toxins that inhibit neurotransmitter release
C;Comment: The heavy chain mediates the binding of toxin to cell receptors while the lig
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin
F;2-422/Product: botulinum neurotoxin type E light chain #status predicted <LIG>
F;423-1251/Product: botulinum neurotoxin type E heavy chain #status predicted <HEA>
F;412-426/Disulfide bonds: #status predicted
Query Match 5.1%; Score 110.5; DB 2; Length 1251;
Best Local Similarity 20.1%; Pred. No. 6.9;
Matches 96; Conservative 83; Mismatches 168; Indels 131; Gaps 27;
QY 1 NILRGESLNKSLPI-----LHEWKFFDYDFGSDERRQDAILSGEYDYKNYPSDI 50
```



301 FYTLPTSLYF--TASATTELQKNNILISVENSEQASLIGNISISF--LEKXTIPD 356  
182 TMVYIADEK-GEGLIVVHNSDDSPHRLTSNTFYDPKFTQMTIDGSGYTAQDGGISGVALS 240  
357 KLFEIITKANPHKIIICSSVLSSFERIKSIQNVETFTNEINKLDE---AKASVINIGII 413  
241 PMTNLY---YSPVASTSL-----YYUNT----- 261  
414 PLNQSFTKEFLYFITSSELEKTYLNTNKKLKNLLELDNLKBEFVVHKDHGIGQFL 473  
262 --BOFTSD-----YQNDIHVEGVONILDTQSSAKVSKSGVLFFGLVGDSA--- 307  
474 KLEAFKIQGLHDFKILYFGNDKLYPVVENI-----EVIKYG-----SDNAELN 519  
308 -LG--CWNEHRTLERHNRITVAQSDETPLQIASMKIEXY---HVPIDFYI-----N 355  
520 KLGSAVWNSKAKLKNRIKEI--SLHLIQIAAKRKLNIPTIELDLEAYDKFCANFPFSE 577  
356 REVILVLNKKQKQKQVNDNFND-----DYVNF 381  
578 TEDQLTAINDIREDLTGLMLDRLICGDVGF 608

RESULT 14  
F82885  
hypothetical protein UUA482 [imported] - Ureaplasma urealyticum  
C;Species: Ureaplasma urealyticum  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C;Accession: F82885  
R;Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.  
submitted to GenBank, February 2000  
A;Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min  
A;Reference number: A82870  
A;Accession: F82885  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-4688 <GLA>  
A;Cross-references: UNIPARC:UPI0000110213; GB:AF222894; NID:96899476; PIDN:  
A;Experimental source: serovar 3; biovar 1  
C;Genetics:  
A;Gene: UUA482  
A;Genetic code: SGC3

Query Match 5.0%; Score 109; DB 2; Length 4688;  
Best Local Similarity 17.6%; Pred. No. 62;  
Matches 87; Conservative 64; Mismatches 169; Indels 174; Gaps 20;  
28 DERQDAILSG-----EYDYKNYPSPDIDQWHDKIFVTMLRYNG----- 66  
1824 DKKEYNLVLNKLKPGRRYSLKKNIKKEVDNGQDHEFVKKEINVNNSFDVNLQSEITASSVE 1883  
67 -----VPSSLNVISKVGDGGLLPQYPMWSFAKYDDCSGIVASAKLAIDKCDRLWLDS 121  
1884 EINRDPKLNQTTIKI-----NLKDDNDILKTDIATIDYNEQKVA 1927  
122 GLVNTQPMCSPLKLTDFLTSSQLLQVEIPHDVAVNATTKGRLSSLAQSLDCNTNSD 181  
1928 IVKNTAQ---NOKYLEAITNLVFNKNVLIK-----KIEFKNLSQTFIKVGNNT 1974  
182 TMVYIADEKGEGLIVYHN-----SDD-----SFRLTSNTFDY--DPKFTK-- 220  
1975 NVIY--DESNLKLINNDQIIGPLSTDVNSTQNVANNKHVISSTLDKFNPHISKNL 2032  
221 -----WTIDGESYTAQDGGISGVALSPMTNNLYSPVASTSLYYVNTTEQ-----FRTSDYQ 270  
2033 KFKLKFQNINGE-----VVTSPILTNSIIIVNNKKNVINTLDMLK 2073  
271 QN-----DIHY-----EGVQVILDTQSSAKVSKSGVLFFGLVGDSA 307  
2074 SNQLYRLVDVYIDNNDNTINDKNKPVANNVTRIIDIPAGKTTISKSN----- 2122  
308 LGCWNEHRTLERHNRITVAQSD-----ETLQIASMK-----I 340

2123 -NTWNTTSSQPEF--VINSDDGNEVLNLEATISFKKGQTLTFVKVNIKONNKYLI 2179  
341 KEAXPHVPFIDRYINREYILVLSNMKQK--WVANDNFDDVNPVPRMANVNNELILNTRCEN 399  
2180 KQQTITLPEPNRYVLSILLAKPNKTKPLVVEILNKLNDISFQTAGNYK--VQIKSQN 2237  
400 PNDRTPFKISIHIL 413  
2238 PSTVDTKQRIKLKL 2251

RESULT 15  
E71619  
RAD2 endonuclease PFB0265c - malaria parasite (Plasmodium falciparum)  
C;Species: Plasmodium falciparum  
C;Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 09-Jul-2004  
C;Accession: E71619  
R;Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;  
; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.;  
Science 282, 1126-1132, 1998  
A;Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.  
A;Reference number: A71600; MUID:99021743; PMID:9804551  
A;Accession: E71619  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-1516 <GAR>  
A;Cross-references: UNIPROT:O96154; UNIPARC:UPI000007D036; GB:AE001383; GB:AE001362; NID  
A;Experimental source: clone 3D7  
C;Genetics:  
A;Gene: PFB0265c

Query Match 4.9%; Score 107.5; DB 2; Length 1516;  
Best Local Similarity 20.6%; Pred. No. 15;  
Matches 84; Conservative 69; Mismatches 141; Indels 113; Gaps 18;  
9 NKSLPILHEWKFFDYDFGSDERRQDAILSGEYDYKNN--YPSDIDQWH--DKIFVTMLR- 63  
848 NKSL-----YEDGENFITRNPITNEYEKNNIIVISDEQKYNEEDIIFKDKIKE 897  
64 --YNGVPSSLNV-----ISKVGGGGLLPQYPMWSFAKYDDCSGIVASAKLAI----- 110  
898 KEXNNDTSSDDFENCVSQEKI-----YVNEKIEEYNNKNDKSSSSSIILEIKY 948  
111 --DKCDLWLVDLSGLVNTQPMCSPLKLTDFLTSSQLLQVEIPHD-----VAVNATTKG 163  
949 KXKKDEL-----VSPNLCV-----LDEFEHSDNLENNYISVSSDDMK 987  
164 GRLLSLAVQSLDCNTNSDITMVIYIADEKGEGLIVVHNSDDSPHRLTSNTFYDPKFTQMTI 223  
988 TNVSKNNTGVKENKVDKTNVEY-DKKGDDGVIEISFEDS-HKLEESKFDDNNNIYDND 1045  
224 DGESYTAQDGGISGVALSPMTNNLYSPVASTSLYYVNTTEQRTSDYQNDIHVEGVQNIL 283  
1046 ELEKNLSKDYISDVOKNHV--NNIY-----NIE--RGEDEREN-----BFVENKI 1086  
284 DTQSSAKVSKSGVLFFGLVGDSALCGWNEHRTLERHNRITVAQSDETLQIASMKIKEA 343  
1087 QSTESHKS-----NEFICTENKSLKQYMSKEDISNVRILKSDD- 1125  
344 XPHVPIFDRYINREYILVLSNMKQKQVNDNFDDVNPVPRMANVNN 390  
1126 -----INNLSKQNYFEILLDKKQVMDNFQNNIENQNDKLKEDKLE 1166

Search completed: May 3, 2006, 19:10:12  
Job time : 35.0175 secs

This Page Blank (usps)

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2006, 18:54:38 ; Search time 191.204 Seconds  
(without alignments)

1523.941 Million cell updates/sec

Title: US-10-525-567-4

Perfect score: 2178

Sequence: 1 NIIRGESLNKSLPILHEWKF.....NTRCENPDNDRTPFKISLH 413

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 05.80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2177	99.9	432	1	MRJPI_APIME
2	2177	99.9	432	2	Q548D6_APIME
3	2023	92.9	433	2	Q8MMJ1_APICE
4	2020	92.7	433	2	Q5VLE2_APICC
5	1454.5	66.8	443	2	Q6IMJ9_APIME
6	1441	66.2	452	1	MRJP2_APIME
7	1411	64.8	463	2	Q8MMJ0_APICE
8	1411	64.8	463	2	Q50H66_APICE
9	1411	64.8	468	2	Q5VK55_APICC
10	1406.5	64.6	544	1	MRJP3_APIME
11	1401	64.3	437	2	Q6W383_APIME
12	1354	62.2	579	2	Q6QN15_APICE
13	1343.5	61.7	598	2	Q5VK56_APICC
14	1300.5	59.7	598	1	MRJP5_APIME
15	1282	58.9	485	2	Q6QN16_APICE
16	1273.5	58.5	416	2	Q6TGR0_APIME
17	1216.5	55.9	464	1	MRJP4_APIME
18	1190	54.6	423	2	Q4ZJX1_APIME
19	1053	48.0	222	2	Q51224_APICE
20	806.5	37.0	222	2	Q51222_APICE
21	781	35.9	220	2	Q51223_APICE
22	508	23.3	463	2	Q9V4C0_DROME
23	492	22.6	406	2	Q7Q4T5_DROSOPHILA
24	466.5	21.4	461	2	Q7PHB5_ANOGA
25	464	21.3	497	2	Q7Z0H1_9DIPT
26	463.5	21.3	540	2	Q86CU0_9DIPT
27	462.5	21.2	541	1	YELL_DROA
28	462.5	21.2	541	2	O8ISG5_DROVI
29	461	21.2	494	2	Q7Z0H2_9DIPT
30	458.5	21.1	546	2	Q5IBN4_9DIPT
31	455.5	20.9	411	2	Q6DLY9_APIME

#### RESULT 1

MRJPI\_APIME STANDARD; PRT; 432 AA.

AC O18330; 20.8 453 462 2 Q7PMG1\_ANOGA  
34 452.5 20.8 541 1 YELL\_DROMA  
35 452.5 20.8 541 1 YELL\_DROME  
36 451.5 20.7 541 1 YELL\_DROSI  
37 450.5 20.7 568 1 YELL\_DROER  
38 448 20.6 433 2 Q7KW97\_DROME  
39 448 20.6 438 2 Q9VJQ3\_DROME  
40 447.5 20.5 568 1 YELL\_DROMD  
41 446.5 20.5 543 2 Q9BI26\_DROAN  
42 445.5 20.5 560 1 YELL\_DROPS  
43 442.5 20.3 432 2 Q9WIR1\_DROME  
44 442.5 20.3 568 1 YELL\_DROSU  
45 435.5 20.0 453 2 Q9VJ15\_DROME

#### ALIGNMENTS

Q7pmg1 anopheles g  
P62408 drosophila  
P09957 drosophila  
P62407 drosophila  
Q9bi23 drosophila  
Q9gp81 drosophila  
Q7kw97 drosophila  
Q9vjq3 drosophila  
Q9gp71 drosophila  
Q9bi26 drosophila  
Q9bi18 drosophila  
Q9wlr1 drosophila  
Q02437 drosophila  
Q9vj15 drosophila

28-FEB-2003 (Rel. 41, Last sequence update)  
10-MAY-2005 (Rel. 47, Last annotation update)  
Major royal jelly protein 1 precursor (MRJP-1) (Bee-milk protein).  
Name=MRJP1;  
OS Apis mellifera (Honeybee).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Hymenoptera; Apoidea; Apoidea;  
OC Apidae; Apis.  
OX NCBI\_TaxID=7460;  
[1]  
NP NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 39-48; 63-71 AND 361-380.  
RP TISSUE=Hypopharyngeal gland;  
RC MEDLINE=98055707; PubMed=9395329;  
RA Ohashi K., Natori S., Kubo T.;  
RT "Change in the mode of gene expression of the hypopharyngeal gland  
cells with an age-dependent role change of the worker honeybee Apis  
mellifera L.";  
RT Eur. J. Biochem. 249:797-802(1997).  
[2]  
RN NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 20-30.  
RP TISSUE=Head;  
RC MEDLINE=99007754; PubMed=9791542; DOI=10.1007/s000180050229;  
RA Schmitzova J., Klaudivny J., Albert S., Schroeder W., Schreckengost W.,  
Hanes J., Judova J., Simuth J.;  
RT "A family of major royal jelly proteins of the honeybee Apis mellifera  
L.";  
RT Cell. Mol. Life Sci. 54:1020-1030(1998).  
RL -!- FUNCTION: May play an important role in honeybee nutrition. Most  
abundant protein found in the royal jelly which is the food of the  
queen honey bee larva. The royal jelly determines the development  
of the young larvae and is responsible for the high reproductive  
ability of the honeybee queen.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Found in the hypopharyngeal glands of the  
worker honeybee.  
CC -!- DEVELOPMENTAL STAGE: Produced in the cephalic glands of both the  
nurse bee and the forager bee. This bee milk protein changes to  
alpha-glucosidase in accordance with the age-dependent role change  
of the worker bee.  
CC -!- SIMILARITY: Belongs to the major royal jelly protein family.  
-----  
This Swiss-Prot entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use as long as its content is in no way modified and this statement is not  
removed.  
EMBL; D79207; BAA23639.1; -; mRNA.  
EMBL; AF000633; AAC61895.1; -; mRNA.

```

DR Ensembl; ENSAPMG00000007331; Apis mellifera.
DR InterPro; IPR009007; Pept_Aspartc_cat.
DR InterPro; IPR003534; Royaljelly.
DR PANTHER; PTHR10009; Royaljelly; 1.
DR pEAM; PF03022; MRJP; 1.
DR PRINTS; PR01366; ROYALJELLY.
KW Direct protein sequencing; Glycoprotein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 432 Major royal jelly protein 1.
FT CARBOHYD 28 24 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 144 144 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 177 177 N-linked (GlcNAc... ) (Potential).
SQ SEQUENCE 432 AA; 48886 MW; 9F42BF08D34A1A7B CRC64;

Query Match 99.9%; Score 2177; DB 1; Length 432;
Best Local Similarity 99.8%; Pred. No. 5.9e-151;
Matches 412; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 NILRGESLNKSLPILHEWKFFDYDFGSDERRQDAILSGEYDYKNNYPSDIDQWHDKIFVT 60
20 NILRGESLNKSLPILHEWKFFDYDFGSDERRQDAILSGEYDYKNNYPSDIDQWHDKIFVT 79

61 MLRYNGVPSSLNVISKVGDDGGLLPYPDPWSFAKYDDCSGIVSASKLAIDKCDRLWVLD 120
80 MLRYNGVPSSLNVISKVGDDGGLLPYPDPWSFAKYDDCSGIVSASKLAIDKCDRLWVLD 139

121 SGLVNNTPQMCSPKLLTFDITTSQLLKQVEIPHDVAVNATTGKRLSSLAQSLDCNTNS 180
140 SGLVNNTPQMCSPKLLTFDITTSQLLKQVEIPHDVAVNATTGKRLSSLAQSLDCNTNS 199

181 DTWVYIADKEGGLIYVHNSDDSFHRLTSNTFDYDPKFTKMTIDGESYTAQDGISGMALS 240
200 DTWVYIADKEGGLIYVHNSDDSFHRLTSNTFDYDPKFTKMTIDGESYTAQDGISGMALS 259

241 PMTNLYYSPVASTSLYYNTEQFRTSDYQONDIHYEGVQNILDTOSSAKVSKGVLPFF 300
260 PMTNLYYSPVASTSLYYNTEQFRTSDYQONDIHYEGVQNILDTOSSAKVSKGVLPFF 319

301 GLVGDSALGCMWNEHRTLERNHIRTVAQSDETLQMIASMKIKEALPHVIPFDRIYINREYIL 360
320 GLVGDSALGCMWNEHRTLERNHIRTVAQSDETLQMIASMKIKEALPHVIPFDRIYINREYIL 379

361 VLSNKKQKQWVNDNFDDVNFRIIMNANVNELIINTRCENPDNDRTPFKISIH 413
380 VLSNKKQKQWVNDNFDDVNFRIIMNANVNELIINTRCENPDNDRTPFKISIH 432

RESULT 3
Q8MMJ1.APICE
ID Q8MMJ1.APICE PRELIMINARY; PRT; 433 AA.
AC Q8MMJ1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Major royal jelly protein MRJP1 precursor (Major royal jelly protein
DE 1).
GN Name=MRJP1; Synonyms=mrjp1;
OS Apis cerana (Indian honeybee).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
OC Apidae; Apis.
OX NCBI_TaxID=7461;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Nurse heads;
RA Sittipraneed S.; Imjongjirak C.;
RT "Molecular Cloning of Major Royal Jelly Protein (MRJP1) cDNA from Apis
RT cerana in Thailand.";
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Imjongjirak C., Klinbunga S., Sittipraneed S.;
RT "Cloning, Expression and Genomic Organization of Genes Encoding Major
RT Royal Jelly Protein 1 and 2 of the Honeybee (Apis cerana).";
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF525776; AAM88281.2; -; mRNA.
DR EMBL; AY515688; AAS88556.1; -; Genomic_DNA.
DR InterPro; IPR003534; Royaljelly.
DR Pfam; PF03022; MRJP; 1.
DR PRINTS; PR01366; ROYALJELLY.
KW Signal.
FT SIGNAL 1 20 Potential.
SQ SEQUENCE 433 AA; 48969 MW; FEC24D6A4BC7BF2F CRC64;

Query Match 92.9%; Score 2023; DB 2; Length 433;
Best Local Similarity 91.3%; Pred. No. 1.1e-139;
Matches 377; Conservative 20; Mismatches 16; Indels 0; Gaps 0;

```

```

QY 1 NILRGESLNKSLPILHEWKFFDYDFGSDERRQDAILSGEYDYKKNYPSDDIDQWHDKLFVT 60
DB 21 SILRGESLNKSLVLEHWKFFDYDFGSDERRQDAILSGEYDYKKNYPSDDVDQWHDKLFVT 80
QY 61 MLRYNGVPSSLNLSKVGCGPLLOPYPDWSFAKYDDCSCGIVSASKLAIDKCDRLWVLD 120
DB 81 MLRYNGVPSSLNLSKVGCGPLLOPYPDWSFAKYDDCSCGIVSASKLAIDKCDRLWVLD 140
QY 121 SGLVNNTQPMCSKPLLTDLTTSQLLKQVEIPHDAVNAATTGKRLSSSLAVQSLDCNTNS 180
DB 141 SGLVNNTQPMCSKPLLTDLTTSQLLKQVEIPHDAVNAATTGKRLSSSLAVQSLDCNTNS 200
QY 181 DTWVIADKEGGLIVYHNSDSEFHLRTSNTFDYDPKFTKMTIDGESYTAQDGISGMALS 240
DB 201 DTWVIADKEGGLIVYHNSDSEFHLRTSNTFDYDPKFTKMTIDGESYTAQDGISGMALS 260
QY 241 PMTNLYYSPVASTSLYYVNTQFRTSDYQONDHYEGVQNIILDTQSSAKVSKGVLF 300
DB 261 PMTNLYYSPVASTSLYYVNTQFRTSDYQONDHYEGVQNIILDTQSSAKVSKGVLF 320
QY 301 GLVGDSALGCWNEHRLTLEHNRIRTVASQDETQMIASMKIKEAXPHVPIFDRIYINREYIL 360
DB 321 GLVGDSALGCWNEHRLTLEHNRIRTVASQDETQMIASMKIKEAXPHVPIFDRIYINREYIL 380
QY 361 VLSNMQKQWNNDFNFDVNFIRIMANVNELILNTRCENPDNDRTPPKISIH 413
DB 381 VLSNMQKQWNNDFNFDVNFIRIMANVNELILNTRCENPDNDRTPPKISIH 433

RESULT 4
Q5VLE2 APICCC
ID Q5VLE2 APICCC PRELIMINARY; PRT; 433 AA.
AC Q5VLE2;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DE Major royal jelly protein MRJ1.
OS Apis cerana cerana (Oriental honeybee).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
OC Apidae; Apis.
OX NCBI_TaxID=94128;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Head;
RA Songkun S., Shenglu C., Stefan A., Boxiong Z.;
RT "Molecular Cloning of MRJ1 cDNA from Apis cerana cerana in China.";
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY279539; AAQ16677.1; -; mRNA.
DR InterPro; IPR003534; Royaljelly.
DR Pfam; PF03022; MRJP; 1.
DR PRINTS; PR01366; ROYALJELLY.
DR SEQUENCE 433 AA; 49045 MW; 86D6C115EF40B9B5 CRC64;

Query Match 92.7%; Score 2020; DB 2; Length 433;
Best Local Similarity 91.3%; Pred. No. 1.9e-139;
Matches 377; Conservative 20; Mismatches 16; Indels 0; Gaps 0;

QY 1 NILRGESLNKSLPILHEWKFFDYDFGSDERRQDAILSGEYDYKKNYPSDDIDQWHDKLFVT 60
DB 21 SILRGESLNKSLVLEHWKFFDYDFGSDERRQDAILSGEYDYKKNYPSDDVDQWHDKLFVT 80
QY 61 MLRYNGVPSSLNLSKVGCGPLLOPYPDWSFAKYDDCSCGIVSASKLAIDKCDRLWVLD 120
DB 81 MLRYNGVPSSLNLSKVGCGPLLOPYPDWSFAKYDDCSCGIVSASKLAIDKCDRLWVLD 140
QY 121 SGLVNNTQPMCSKPLLTDLTTSQLLKQVEIPHDAVNAATTGKRLSSSLAVQSLDCNTNS 180
DB 141 SGLVNNTQPMCSKPLLTDLTTSQLLKQVEIPHDAVNAATTGKRLSSSLAVQSLDCNTNS 200
QY 181 DTWVIADKEGGLIVYHNSDSEFHLRTSNTFDYDPKFTKMTIDGESYTAQDGISGMALS 240
DB 201 DTWVIADKEGGLIVYHNSDSEFHLRTSNTFDYDPKFTKMTIDGESYTAQDGISGMALS 260
QY 241 PMTNLYYSPVASTSLYYVNTQFRTSDYQONDHYEGVQNIILDTQSSAKVSKGVLF 300
DB 261 PMTNLYYSPVASTSLYYVNTQFRTSDYQONDHYEGVQNIILDTQSSAKVSKGVLF 320
QY 301 GLVGDSALGCWNEHRLTLEHNRIRTVASQDETQMIASMKIKEAXPHVPIFDRIYINREYIL 360
DB 321 GLVGDSALGCWNEHRLTLEHNRIRTVASQDETQMIASMKIKEAXPHVPIFDRIYINREYIL 380
QY 361 VLSNMQKQWNNDFNFDVNFIRIMANVNELILNTRCENPDNDRTPPKISIH 413
DB 381 VLSNMQKQWNNDFNFDVNFIRIMANVNELILNTRCENPDNDRTPPKISIH 433

RESULT 5
Q6IMJ9 APIME
ID Q6IMJ9 APIME PRELIMINARY; PRT; 443 AA.
AC Q6IMJ9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Major royal jelly protein 7.
GN Name=MRJP7;
OS Apis mellifera (Honeybee).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
OC Apidae; Apis.
OX NCBI_TaxID=7460;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC PubMed=15037093; DOI=10.1016/j.jinsphys.2003.09.008;
RA Albert S., Klaudiny J.;
RT "The MRJP/YELLOW protein family of Apis mellifera: identification of
RT new members in the EST library.";
RL J. Insect Physiol. 50:51-59(2004).
CC -I- MISCCELLANEOUS: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ third party annotation (TPA) entry.
DR EMBL; BK001420; DAA01512.1; -; mRNA.
DR Ensembl; ENSAPMG00000007331; Apis mellifera.
DR InterPro; IPR003534; Royaljelly.
DR Pfam; PF03022; MRJP; 1.
DR PRINTS; PR01366; ROYALJELLY.
DR SEQUENCE 443 AA; 50541 MW; 8916272B4A4DB421 CRC64;

Query Match 66.8%; Score 1454.5; DB 2; Length 443;
Best Local Similarity 65.5%; Pred. No. 5e-98;
Matches 270; Conservative 66; Mismatches 63; Indels 13; Gaps 4;

QY 2 ILRGES---LNKSLPILHEWKFFDYDFGSDERRQDAILSGEYDYKKNYPSDDIDQWHDKIF 58
DB 19 ILRENSARNLKNLSKLVHEWKYIDYDFGSEBKRAALQSDEYDHTKKNYPPFDQWDRKTF 78
QY 59 VTMLRYNGVPSSLNLSKVGCGPLLOPYPDWSFAKYDDCSCGIVSASKLAIDKCDRLWV 118
DB 79 VTMLRYNGVPSSLNLSKVGCGPLLOPYPDWSFAKYDDCSCGIVSASKLAIDKCDRLWV 138
QY 119 LDSGLVNNTQPMCSKPLLTDLTTSQLLKQVEIPHDAVNAATTGKRLSSSLAVQSLDCNT 178
DB 139 LDSGLVNNTQPMCSKPLLTDLTTSQLLKQVEIPHDAVNAATTGKRLSSSLAVQSLDCNT 195
QY 179 NSDTMVIADKEGGLIVYHNSDSEFHLRTSNTFDYDPKFTKMTIDGESYTAQDGISGM 238
DB 196 SVNTLVYIADNKGDLIVYQNSDSEFHLRTSNTFDYDPKFTKMTIDGESYTAQDGISGM 255
QY 239 LSPMTNNLYYSPVASTSLYYVNTQFRTSDYQONDHYEGVQNIILDTQSSAKVSKGV 298
DB 256 LSPMTNNLYYSPVASTSLYYVNTQFRTSDYQONDHYEGVQNIILDTQSSAKVSKGV 315
QY 299 PFGLVGDPSALGCWNEHRLTLEHNRIRTVASQDETQMIASMKIKEAXPHVPI 356
DB 316 PFGLVGDPSALGCWNEHRLTLEHNRIRTVASQDETQMIASMKIKEAXPHVPI 375

```









RT protein\_RJP571; longer repetitive region on C-terminus and homology to  
RT Drosophila melanogaster yellow protein.";  
RL J. Apicultural Res. 35:63-68(1996).  
RN [3]  
RN NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 21-36.  
RC TISSUE=Head;  
RX MEDLINE=99007754; PubMed=9791542; DOI=10.1007/s000180050229;  
RA Schmitzova J., Klaudiny J., Albert S., Schroeder W., Schreckengost W.,  
RA Hanes J., Judova J., Simuth J.;  
RT "A family of major royal jelly proteins of the honeybee Apis mellifera  
L.";  
RL Cell. Mol. Life Sci. 54:1020-1030(1998).  
RN [4]  
RN NUCLEOTIDE SEQUENCE, AND PARTIAL PROTEIN SEQUENCE OF 45-59; 70-77;  
RP 106-120; 214-224 AND 243-253.  
RC TISSUE=Hypopharyngeal gland;  
RX MEDLINE=98055707; PubMed=9395329;  
RA Chashi K., Natori S., Kubo T.;  
RT "Change in the mode of gene expression of the hypopharyngeal gland  
RT cells with an age-dependent role change of the worker honeybee Apis  
mellifera L.";  
RL Eur. J. Biochem. 249:797-802(1997).  
CC -!- FUNCTION: May play an important role in honeybee nutrition. It is  
found in the royal jelly which is the food of the queen honey bee  
larva. The royal jelly determines the development of the young  
larvae and is responsible for the high reproductive ability of the  
honeybee queen.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Hypopharyngeal glands of nurse honey bees.  
CC -!- DEVELOPMENTAL STAGE: Produced by the cephalic glandular system of  
the nurse honey bee.  
CC -!- SIMILARITY: Belongs to the major royal jelly protein family.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
DR EMBL; Z26318; CAA81227.1; -; mRNA.  
DR Ensembl; ENSAPMG0000007331; Apis mellifera.  
DR InterPro; IPR003534; Royaljelly.  
DR FANTHER; FTHR10009; Royaljelly.  
DR Pfam; PF03022; MRJP, 1.  
DR PRINTS; PR01366; ROYALJELLY.  
KW Direct protein sequencing; Glycoprotein; Repeat; Signal.  
FT SIGNAL 1 20  
FT CHAIN 21 544 Major royal jelly protein 3.  
FT REPEAT 424 428 1.  
FT REPEAT 429 433 2.  
FT REPEAT 434 438 3.  
FT REPEAT 439 443 4.  
FT REPEAT 444 448 5.  
FT REPEAT 449 453 6.  
FT REPEAT 454 458 7.  
FT REPEAT 459 463 8.  
FT REPEAT 464 468 9.  
FT REPEAT 469 473 10.  
FT REPEAT 474 478 11.  
FT REPEAT 479 483 12.  
FT REPEAT 484 488 13.  
FT REPEAT 489 493 14.  
FT REPEAT 494 498 15.  
FT REPEAT 499 503 16.  
FT REPEAT 504 508 17.  
FT REPEAT 509 513 18.  
FT REPEAT 514 518 19.  
FT REPEAT 519 523 20.  
FT REGION 424 523 20 x 5 AA tandem repeats of [NKR]-Q-N-  
[AGD]-[DNG].  
FT CARBOHYD 183 183 N-linked (GlcNAc...) (Potential).  
FT SEQUENCE 544 AA; 61662 MW; 4C8FFCBA2759F52 CRC64;

Query Match 64.6%; Score 1406.5; DB 1; Length 544;  
Best Local Similarity 64.8%; Pred. NO. 2.2e-94;  
Matches 263; Conservative 62; Mismatches 68; Indels 13; Gaps 2;  
  
QY 7 SLNKSPLILHEWKFPDYDFGSDERRQDAILSGEYDYKNNYPSDIDQWHDKIFVTMLRYNG 66  
DB 32 NLAHSKVIYEWKHDIDFGSDERRDAAIKSEFDHTKYPFPDVRDRDKTFTTIERNG 91  
QY 67 VPSLNVISKVGDGGLLPQYDWSFAKYDCSGIVSASKLAIDKCDRLWLDLSGLVNN 126  
DB 92 VPSLNVVTKKGGKGGPFLRPYDWSFAKYDCSGIVSFAKIAVDKFDRLWLDLSGLVNN 151  
QY 127 TQPMCSFKLLTFDLTTSQLLKQVEIPHDVAVNATTKGRLSSLAVQSLODNTNMTWYI 186  
DB 152 NOPMCSFKLLTFDLTKTSLKVQVEIPHNIAVNATTGMGELVSLAVQAID---RTNTMVYI 208  
QY 187 ADEKGEGLIVYHNSDDSFHRLTSNTFDYDPKFTMTIDGESYTAODGISGMALSPMTNNL 246  
DB 209 ADEKGEGLIYQNSDDSFHRLTSNTFDYDPYFKLTVAGESFTVKNGIYGIASPTVNNL 268  
QY 247 YVSPVASTSLYYVNTQFRTSDYQQNDIHYEGVQNILDTQSSAKVSVKSGVLFGLVGDS 306  
DB 269 YVSPLLSHGLYYVDTQFSNPQVEENNVQYEGSQDILNTQSGKVVSKXGVLFLGLVGS 328  
QY 307 ALGCWNEHRTLEHNIPTVAQSDETLOMTASKIEXAPHVPIDFYINREYILVLSNKM 366  
DB 329 GIACVNEHQVLQRESFDVVAQNEETLOMTIVSMKIMENLPQSGRINDPREGNEYMLALSNNRM 388  
QY 367 QKVVNDNFDDVNFPRIMANVNELILNTRC-----ENPDN 402  
DB 389 QKIINDNFNDVNFRLGAVNDDLMNTRCGRYHNQAGNQADN 434  
  
RESULT 11  
Q6W3E3 APIME  
ID Q6W3E3 APIME PRELIMINARY; PRT; 437 AA.  
AC Q6W3E3;  
DT 05-JUL-2004 (TREMBlrel. 27, Created)  
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
DE Major royal jelly protein MRJP6.  
OS Apis mellifera (Honeybee)  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;  
OC Apidae; Apis.  
OX NCBI\_TaxID=7460;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX PubMed=15037093; DOI=10.1016/j.jinphys.2003.09.008;  
RA Albert S., Klaudiny J.;  
RT "The MRJP/YELLOW protein family of Apis mellifera: identification of  
RT new members in the EST library.";  
RL J. Insect Physiol. 50:51-59(2004).  
DR EMBL; AY113893; AAQ82184.1; -; mRNA.  
DR Ensembl; ENSAPMG00000007331; Apis mellifera.  
DR InterPro; IPR003534; Royaljelly.  
DR Pfam; PF03022; MRJP, 1.  
DR PRINTS; PR01366; ROYALJELLY.  
SQ SEQUENCE 437 AA; 49786 MW; 1B654BB7A7F0B56D CRC64;  
  
Query Match 64.3%; Score 1401; DB 2; Length 437;  
Best Local Similarity 63.3%; Pred. NO. 4.1e-94;  
Matches 260; Conservative 69; Mismatches 76; Indels 6; Gaps 2;  
  
QY 6 ESLNKSPLILHEWKFPDYDFGSDERRQDAILSGEYDYKNNYPSDIDQWHDKIFVTMLRYN 65  
DB 29 KMLEHSNVVHEWKYIDYDFGSDERKQAAIQSEYDYTKNYPFPDQWKNKTFLAVRYD 88  
QY 66 GVPSSLNVISKVGDGGLLPQYDWSFAKYDDCSGIVSASKLAIDKCDRLWLDLSGLV 125  
DB 89 GVPSSLNVISEKINGGCLLPQYDWSWANYKDCSGIVSAYKIAIDKFDRLWLDLSGLIN 148  
QY 126 NTQPMCSFKLLTFDLTTSQLLKQVEIPHDVAVNATTKGRLSSLAVQSLODNTNMTWY 185







File Page Blank (asp.c)

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2006, 19:09:24 ; Search time 44.6142 Seconds  
(without alignments)  
765.341 Million cell updates/sec

Title: US-10-525-567-4

Perfect score: 2178

Sequence: 1 NIIRGSLNKLPIHWEKFP.....NTRCENPDNDRTPFKLSIHL 413

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/1/iaa/5 COMB pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/6 COMB pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/H COMB pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/PTCUS COMB pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/RE COMB pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1 pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	233	10.7	162	2	US-09-270-767-33773
2	233	10.7	162	2	US-09-270-767-48990
3	126	5.8	163	2	US-09-270-767-32344
4	126	5.8	163	2	US-09-270-767-47561
5	116.5	5.3	161	2	US-09-270-767-43507
6	108.5	5.0	1427	2	US-09-902-540-10612
7	105	4.8	159	2	US-09-270-767-32059
8	105	4.8	159	2	US-09-270-767-47276
9	104	4.8	956	2	US-09-134-001C-4452
10	103	4.7	3852	2	US-10-025-225-4
11	103	4.7	4585	2	US-10-025-225-6
12	103	4.7	4588	2	US-10-025-225-8
13	103	4.7	4589	2	US-10-025-225-2
14	102	4.7	471	2	US-09-134-001C-3150
15	101.5	4.7	789	2	US-09-134-000C-4939
16	100.5	4.6	2308	1	US-08-015-973-1
17	100.5	4.6	2308	1	US-08-448-164-1
18	100.5	4.6	2308	2	US-08-081-929-2
19	100.5	4.6	2308	2	US-10-000-954-2
20	100.5	4.6	2314	2	US-09-816-703A-2
21	99	4.5	823	2	US-10-101-464A-923
22	98.5	4.5	461	2	US-09-252-991A-29720
23	98.5	4.5	970	1	US-08-673-789-7
24	98.5	4.5	1007	2	US-08-961-083-216
25	98.5	4.5	1007	2	US-09-536-784-216
26	98.5	4.5	1007	2	US-09-765-271-216
27	98.5	4.5	1007	2	US-09-765-272A-216

28	98.5	4.5	2026	2	US-09-487-558B-86	Sequence 86, Appl
29	97.5	4.5	371	2	US-09-991-181-171	Sequence 171, App
30	97.5	4.5	371	2	US-09-990-444-171	Sequence 171, App
31	97.5	4.5	371	2	US-09-997-333-171	Sequence 171, App
32	97.5	4.5	371	2	US-09-992-598-171	Sequence 171, App
33	97.5	4.5	392	2	US-09-248-796A-16779	Sequence 16779, A
34	97.5	4.5	881	2	US-09-248-796A-18627	Sequence 18627, A
35	97.5	4.5	1608	2	US-09-568-407-1	Sequence 1, Appli
36	96	4.4	432	2	US-09-270-767-46441	Sequence 46441, A
37	96	4.4	650	2	US-09-487-558B-430	Sequence 4304, A
38	95.5	4.4	356	2	US-09-270-767-43404	Sequence 20509, A
39	95	4.4	576	2	US-09-248-796A-20509	Sequence 353, App
40	95	4.4	781	2	US-09-538-092-363	Sequence 253, App
41	94.5	4.3	194	2	US-09-673-395A-253	Sequence 6185, Ap
42	94.5	4.3	669	2	US-09-134-000C-6185	Sequence 5, Appli
43	94.5	4.3	872	1	US-08-387-942C-5	Sequence 2, Appli
44	94.5	4.3	2391	1	US-08-446-855A-2	Sequence 2, Appli
45	94.5	4.3	2391	2	US-09-150-741-2	Sequence 2, Appli

#### ALIGNMENTS

##### RESULT 1

US-09-270-767-33773  
; Sequence 33773, Application US/09270767  
; Patent No. 6703491

; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: Patent In Ver. 2.0

; SEQ ID NO 33773

; LENGTH: 162

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

US-09-270-767-33773

Query Match 10.7%; Score 233; DB 2; Length 162;

Best Local Similarity 42.1%; Pred. No. 8.5e-15;

Matches 48; Conservative 25; Mismatches 29; Indels 12; Gaps 5;

Qy 43 KNTVPSIDQWHDKIFVTMLRY-NGVPSLLNVI---SKKVGSGGLLPQYDMSF---AK 95

Db 11 KNNLPLGIDVHNNRLFVTTPRKNGVPASLTGLPFPK--GVKGPALKPYNWEAHGNPN 68

Qy 96 YDCSGIVSASKLAIDKCDRLWLDLSDGLVN---NTQPMCSPKLTLTDTTSLQL 146

Db 69 NPDCKLMSVYRTAVDRCDRIWLIDSGIVNATINLNCICPKIVVYDLKSDLEI 122

##### RESULT 2

US-09-270-767-48990

; Sequence 48990, Application US/09270767

; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: Patent In Ver. 2.0

; SEQ ID NO 48990

; LENGTH: 162

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

US-09-270-767-48990

Query Match 10.7%; Score 233; DB 2; Length 162;









Db 2063 LRVARVVVRVNI--EDINDNSPVFVGLPYAAVQVDAEPGTLIVQVTAIDK-DKGPNGEV 2119  
Qy 116 LWVL-----DSGLVNTQPMCS-----KLTFDLTTSQLLKQVEIPHDVA 156  
Db 2120 TVVLQDDYGHFEINPNSGNVILKEAFNSDLNIEYGVITLAKDGGKPSLTSVELPITIV 2179  
Qy 157 VNATTGKRLSSLAVALDQNTNSDTM-VYIADEKGBGLVYHNSDDSFHRLTSNTFYD 215  
Db 2180 NKAMPVDPKPPYATASVNDIRMTPILSINATSPGQGIYIIDGDPFKQF---NIDFD 2236  
Qy 216 PKFTMT--IDGE-----SYTAQDGISG-----MALSPMTNN--LYSGFVASTSL 256  
Db 2237 TGVLVKVSPLDYEVTSAYKLTIRASDALTCARAEVTVDLLVNDVNDNPPIFDQPTNTTL 2296  
Qy 257 ---YVNTGQPT-----SDVQOND-IHYEGVQNILDTQSSAKVVSQVGLVFLGVDGSA 308  
Db 2297 SEASLIGTPVLQVVSIDADSENKVMHYQIVQDYNSTDYFHDSSGLILTARMLDHEL 2356  
Qy 309 GCWNEHRTLERNHIRTVAQSDTL--QMIASMKIKEAXPHVIPDRYINREYILVLSNK- 365  
Db 2357 ---VQCTLK---VRSDSGFPLSSEVLVHIYISDVNDNPPVFNQLIYSEYSELAPRG 2410  
Qy 366 -----MOKMVNDFNFDVFRIMNANVNELLILNTRCENPDNDRTPF 407  
Db 2411 HFVTCVQASDADSDFDRLSEYLSG-----NDRTSF 2442

RESULT 13  
US-10-025-225-2  
; Sequence 2, Application US/10025225  
; Patent No. 6852844  
; GENERAL INFORMATION:  
; APPLICANT: Yu, Xuanchuan  
; APPLICANT: Miranda, Miracar  
; APPLICANT: Turner, C. Alexander Jr.  
; TITLE OF INVENTION: No. 6852844el Human Protocadherin Proteins and Polynucleotides En  
; TITLE OF INVENTION: Same  
; FILE REFERENCE: LEX-0285-USA  
; CURRENT APPLICATION NUMBER: US/10/025,225  
; CURRENT FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: 2001-12-19  
; PRIOR FILING DATE: 2000-12-20  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 4589  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-10-025-225-2

Query Match 4.7%; Score 103; DB 2; Length 4589;  
Best Local Similarity 19.4%; Pred. No. 14;  
Matches 79; Conservative 78; Mismatches 162; Indels 88; Gaps 19;  
Qy 62 LRYNGVPSLNVISKVGGGGLLQPYDPWFSKDYDCSGIVSASKLAIDKCDR----- 115  
Db 2063 LRVARVVVRVNI--EDINDNSPVFVGLPYAAVQVDAEPGTLIVQVTAIDK-DKGPNGEV 2119  
Qy 116 LWVL-----DSGLVNTQPMCS-----KLTFDLTTSQLLKQVEIPHDVA 156  
Db 2120 TVVLQDDYGHFEINPNSGNVILKEAFNSDLNIEYGVITLAKDGGKPSLTSVELPITIV 2179  
Qy 157 VNATTGKRLSSLAVALDQNTNSDTM-VYIADEKGBGLVYHNSDDSFHRLTSNTFYD 215  
Db 2180 NKAMPVDPKPPYATASVNDIRMTPILSINATSPGQGIYIIDGDPFKQF---NIDFD 2236  
Qy 216 PKFTMT--IDGE-----SYTAQDGISG-----MALSPMTNN--LYSGFVASTSL 256  
Db 2237 TGVLVKVSPLDYEVTSAYKLTIRASDALTCARAEVTVDLLVNDVNDNPPIFDQPTNTTL 2296  
Qy 257 ---YVNTGQPT-----SDVQOND-IHYEGVQNILDTQSSAKVVSQVGLVFLGVDGSA 308  
Db 2297 SEASLIGTPVLQVVSIDADSENKVMHYQIVQDYNSTDYFHDSSGLILTARMLDHEL 2356

Qy 309 GCWNEHRTLERNHIRTVAQSDTL--QMIASMKIKEAXPHVIPDRYINREYILVLSNK- 365  
Db 2357 ---VQCTLK---VRSDSGFPLSSEVLVHIYISDVNDNPPVFNQLIYSEYSELAPRG 2410  
Qy 366 -----MOKMVNDFNFDVFRIMNANVNELLILNTRCENPDNDRTPF 407  
Db 2411 HFVTCVQASDADSDFDRLSEYLSG-----NDRTSF 2442

RESULT 14  
US-09-134-001C-3150  
; Sequence 3150, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 3150  
; LENGTH: 471  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3150

Query Match 4.7%; Score 102; DB 2; Length 471;  
Best Local Similarity 20.8%; Pred. No. 0.4;  
Matches 64; Conservative 38; Mismatches 112; Indels 94; Gaps 11;  
Qy 99 CSGIVSASKLAIDKCDRLVLDGLVNTQPMCSPKLLTF-----DLTTSQLLKQVEIPHD 154  
Db 13 CDSMIKEQKLKYK-----WMLITLITFILLCLIIIFPKDTRLASSEIDEAERSND 67  
Qy 155 VAVNATTGKRLSSLAVALDQNTNSDTMVIADKEGGLVYHNSDDSFHRLTSNTFYD 214  
Db 68 IA-NLPHSKS--LSDISALDLNASLENFQELIYDQKGRKLIQTSN-----DNTLAY 116  
Qy 215 DPKFTMTIDGESYTAQDGISGSMALSPMTNNLYSPVASTSLYVNTGQPTSDYQQNDI 274  
Db 117 DNKIDFKHPIRIHQSHGIN-----YLVITEPIRSKDFSGYSV 155  
Qy 275 HYGCVQNILDTQSSAKVVSQVGLVFLGVDGSAALGCWNEHRTLERNHIRTVAQSDTLQ 334  
Db 156 LVHSLQVYDNLVKSLEYVA---LAFGLIA-----TIITAGVS----- 189  
Qy 335 IASMKIKEAXPHVIPDRYINREYILVLSNKQKQKQVNDNFDDVFRIMNANVNELLILN 394  
Db 190 -----YIFSSQITKP-IVTMSNKQKQIRRDGQ-----NKLELT 222  
Qy 395 TRCENPDN 402  
Db 223 TNYEETDN 230

RESULT 15  
US-09-134-000C-4939  
; Sequence 4939, Application US/09134000C  
; Patent No. 6617156  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 032796-032  
; CURRENT APPLICATION NUMBER: US/09/134,000C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/055,778



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2006, 19:10:39 ; Search time 144.89 Seconds  
(without alignments)  
1190.997 Million cell updates/sec

Title: US-10-525-567-4  
Perfect score: 2178  
Sequence: 1 NLRGSLNKLPLHWEKF.....NTRCENPDNDRTPPKISHL 413  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues  
Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA\_Main:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	452.5	20.8	541	6	US-11-097-143-7173
2	448	20.6	438	6	US-11-097-143-7005
3	435.5	20.0	453	6	US-11-097-143-40164
4	422	19.4	426	6	US-11-097-143-25233
5	415	19.1	412	6	US-11-097-143-20991
6	411.5	18.9	415	6	US-11-097-143-20973
7	366	16.8	429	6	US-11-097-143-28038
8	358	16.4	429	6	US-11-097-143-42198
9	357	16.4	379	6	US-11-097-143-41712
10	343.5	15.8	409	6	US-11-097-143-25236
11	331.5	15.2	316	6	US-11-097-143-1674
12	258	11.8	393	6	US-11-097-143-699
13	233	10.7	382	6	US-11-097-143-32673
14	229	10.5	394	4	US-10-481-180-17
15	209.5	9.6	400	4	US-10-481-180-18
16	177	8.1	388	4	US-10-156-761-7699
17	156	7.2	342	6	US-11-097-143-16716
18	122.5	5.6	575	4	US-10-724-972A-7483
19	116	5.3	946	4	US-10-032-585-7900
20	112.5	5.2	1252	4	US-10-354-774-52
21	112.5	5.2	1252	4	US-10-271-012-52
22	112.5	5.2	1252	4	US-10-452-024-128
23	112.5	5.2	1252	4	US-10-729-122-52
24	112.5	5.2	1252	4	US-10-729-039-52
25	112.5	5.2	1252	5	US-10-729-527-52
26	112.5	5.2	1252	5	US-10-727-898-52
27	112.5	5.2	1252	5	US-10-728-696-52

28	112.5	5.2	1252	6	US-11-001-241-52	Sequence 52, Appl
29	111	5.1	4643	6	US-11-097-143-41619	Sequence 41619, A
30	110.5	5.1	1250	4	US-10-354-774-50	Sequence 50, Appl
31	110.5	5.1	1250	4	US-10-271-012-50	Sequence 50, Appl
32	110.5	5.1	1250	4	US-10-729-122-50	Sequence 50, Appl
33	110.5	5.1	1250	4	US-10-729-039-50	Sequence 50, Appl
34	110.5	5.1	1250	5	US-10-729-527-50	Sequence 50, Appl
35	110.5	5.1	1250	5	US-10-727-898-50	Sequence 50, Appl
36	110.5	5.1	1250	5	US-10-728-696-50	Sequence 50, Appl
37	110.5	5.1	1250	6	US-11-001-241-50	Sequence 5, Appl
38	110.5	5.1	1251	4	US-10-452-024-5	Sequence 124, App
39	110.5	5.1	1251	4	US-10-452-024-124	Sequence 129, App
40	110.5	5.1	1251	4	US-10-452-024-129	Sequence 125, App
41	110.5	5.1	1252	4	US-10-452-024-125	Sequence 10, Appl
42	110.5	5.1	1252	4	US-10-205-516-10	Sequence 126, App
43	110.5	5.1	1255	4	US-10-452-024-126	Sequence 71379, A
44	110	5.1	391	4	US-10-282-122A-71379	Sequence 130, App
45	110	5.1	742	5	US-10-745-237-130	

ALIGNMENTS

RESULT 1  
US-11-097-143-7173  
; Sequence 7173, Application US/11097143  
; Publication No. US20050208558A1  
; GENERAL INFORMATION:  
; APPLICANT: Venter, J. Craig  
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
; TITLE OF INVENTION: DROSOPHILA GENES.  
; FILE REFERENCE: CL000728  
; CURRENT APPLICATION NUMBER: US/11/097,143  
; CURRENT FILING DATE: 2005-04-04  
; PRIOR APPLICATION NUMBER: 60/157,832  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: 60/160,191  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR APPLICATION NUMBER: 60/161,932  
; PRIOR FILING DATE: 1999-10-28  
; PRIOR APPLICATION NUMBER: 60/164,769  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR APPLICATION NUMBER: 60/173,383  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: 60/175,693  
; PRIOR FILING DATE: 2000-01-12  
; PRIOR APPLICATION NUMBER: 60/184,831  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/191,637  
; PRIOR FILING DATE: 2000-03-23  
; NUMBER OF SEQ ID NOS: 43008  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7173  
; LENGTH: 541  
; TYPE: PRT  
; ORGANISM: DROSOPHILA

Query Match	20.8%;	Score 452.5;	DB 6;	Length 541;
Best Local Similarity	30.3%;	Pred. No. 2.5e-31;		
Matches	125;	Conservative	90;	Mismatches 150;
			Indels	47;
			Gaps	17;
Qy	16	HEWKFFDYFGSDRRQDAILSGEYDYKNYPSIDQWHDKIFVTMLRY-NGVPSSLNVI	74	
Db	29	YSWSQLDFAFENTRLKQALASGDYIFQNALPVGHEFNRLFTVTPWRDGPATLYI	88	
Qy	75	S-KKVGCGPLQYPPDWSFAKYDDC-SGIVSASKLAIDKCDRLWLDSGLV-NTNTP	129	
Db	89	NMDRSLTGSPELIPYDPMWRSNTAGDCAISITATRIKVDCEGRWLWLDGTGIGNTTN	148	
Qy	130	MCSPKLLTFDLTTSQLLKQVEIPHDVAVNATTGKRLSSLAQSLDCNTN-SDTWVYIAD	188	

Db 149 PCYAVNVFLDTTDIRRYELP-GVDTNPNT---FIANIAV---DIGKNCDDAYAFAD 201  
QY 189 EKEGGLIVYHNSDDSPHRLTSNTFDY-DPKFTKMTIDGESYT-AQDGISGMALSPMTNN- 245  
Db 202 ELGGLIAYSWELNKSXWRFSAHSYFFPDPLRGDFNAGINFQWGEIGFQMSLSPRSBG 261  
QY 246 ---LYYSPVASTSLYYVNT---EQFRTSYQONDHYEGVQNILDPQSSAKVYKSGVL 298  
Db 262 YRTLYSPPLASHRQFAVSTRILRDTRTSDSYHDFVALD--ERGPSHTTSRVMSDDGIE 319  
QY 299 FFLGVGDSALGCWNEHRTL---ERNIRTVQASDETLQMIASMKIEXAPHVPIPDRIY 354  
Db 320 LFNLDIQNAVGCW--HSSMYSPOFHGI--VDRDDVGLVFPADVKID----- 363  
QY 355 NREYILVLSNMKQVNNDFNDFVFRIMNANVNELINTRCENPDNDRT 406  
Db 364 -KNVWVLSDRMPVFLSLDLDYSDTNFRIYTAPLATIENTVCDLRNNAVGP 414

## RESULT 2

US-11-097-143-7005  
; Sequence 7005, Application US/11097143  
; Publication No. US20050208558A1  
; GENERAL INFORMATION:  
; APPLICANT: Venter, J. Craig  
; APPLICANT: et al.  
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
; TITLE OF INVENTION: DROSOPHILA GENES.  
; FILE REFERENCE: CL000728  
; CURRENT APPLICATION NUMBER: US/11/097,143  
; CURRENT FILING DATE: 2005-04-04  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR FILING DATE: 1999-10-28  
; PRIOR FILING DATE: 1999-10-28  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR FILING DATE: 2000-01-12  
; PRIOR FILING DATE: 2000-01-12  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR FILING DATE: 2000-03-23  
; NUMBER OF SEQ ID NOS: 43008  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7005  
; LENGTH: 438  
; TYPE: PRT  
; ORGANISM: DROSOPHILA  
US-11-097-143-7005

Query Match 20.6%; Score 448; DB 6; Length 438;  
Best Local Similarity 28.8%; Pred. No. 4.6e-31;  
Matches 123; Conservative 67; Mismatches 169; Indels 68; Gaps 10;  
QY 18 WKFDYDFGSDERRQDAILSGEYDYKKNYPSDIDQWHDKIFVTMLRYN-GVPSLSNVISK 76  
Db 33 WKQAFDWPTEAEAEAKSNHGYIVENNPLGVVERWQNRIFVTVPWKAGVAATLNYIDI 92  
QY 77 KVGGGGLLOPYDWSFAK-----YDCSGIYSASK 107  
Db 93 NSTEKSPLHYPYSWEANKLPIDVQPDQKTPSGRLDADKAQDAGIQLKDNSTVISTFR 152  
QY 108 LAIDKCDRLWLDVSGVNNTPQMSCKP-----LLTFDLTTSOLLKQVEIPHVDVAVNATT 161  
Db 153 IQVDVCDRLWLDVLTGLAD---ILGSPKQITPNSILVFDLTKDTLLRFTTPAD-----QT 204

QY 162 GKRLSSLAQSLDCNTNSDTPWYIADEKGEGLIVYHNSDDSPHRLTSNTFDYDPKFTKM 221  
Db 205 KEDSFFANIVVDADRSECQDAFAYIPDLGAYGIVYVSLRNDKSYRVKHNFFHFDPLHGDF 264  
QY 222 TIDGESYTAQDGISGMALSPM---TNNLYSPVASTSLYYVN-----TEQFRTSYQON 272  
Db 265 NVGCVNFQWTDGVFLGAVGEMNPDHSDYIFHALASTKEFKVNRVQLONESHTVAGDSYY 324  
QY 273 DIHYEGVQNILDPQSSAKVVS-KSGVLFFGLVGDSSALGCWNEHRTLERNHRTVAOSDET 331  
Db 325 DKYVVGDRG-MNGQSTAEVFDPEGTGIVFTQVNMKDAICWNIKRPYTPDQGLIDSDSHT 383  
QY 332 LQMIASMKIEXAPHVPIPDRIYNREYILVLSNMKQVNNDFNDFVFRIMNANVNEL 391  
Db 384 LVFPNDMKID-----NEGTVIWLSDKMTPLYKELDPDSAVNRYILMGQNRDL 430  
QY 392 ILNTRCE 398  
Db 431 IKGTPCE 437

RESULT 3  
US-11-097-143-40164  
; Sequence 40164, Application US/11097143  
; Publication No. US20050208558A1  
; GENERAL INFORMATION:  
; APPLICANT: Venter, J. Craig  
; APPLICANT: et al.  
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
; TITLE OF INVENTION: DROSOPHILA GENES.  
; FILE REFERENCE: CL000728  
; CURRENT APPLICATION NUMBER: US/11/097,143  
; CURRENT FILING DATE: 2005-04-04  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR FILING DATE: 1999-10-28  
; PRIOR FILING DATE: 1999-10-28  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR FILING DATE: 2000-01-12  
; PRIOR FILING DATE: 2000-01-12  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR FILING DATE: 2000-03-23  
; NUMBER OF SEQ ID NOS: 43008  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 40164  
; LENGTH: 453  
; TYPE: PRT  
; ORGANISM: DROSOPHILA  
US-11-097-143-40164

Query Match 20.0%; Score 435.5; DB 6; Length 453;  
Best Local Similarity 28.4%; Pred. No. 6.3e-30;  
Matches 116; Conservative 85; Mismatches 165; Indels 43; Gaps 12;  
QY 9 NKSLLPLIHEWKFPDYDFGSDERRQDAILSGEYDYKKNYPSDIDQWHDKIFVTMLRY-NGV 67  
Db 19 NDNLRVAYEWREMDFKYANPDQWSAIERGEFEKPANVIPPFGLEVAGHRLFVTLPRMRDGV 78  
QY 68 PSS-----LNVISKVGDGGLLOPYDWSFAKYDDCS-CIVSASKLAIDKCDRLWLD- 120  
Db 79 PASLAYLDLNDTSSK---GPALKPFPFSSWOAHNLQEAPELVPFVRADRCGRWLWDS 134  
QY 121 --SGLVNNTPQMSCKPILLTFDLTTSOLLKQVEIPHVDVAVNATTCKGRSLSLAQSLDCNT 178  
Db 135 RISGVLEQTKIYGAQQLVYVLDLNDLLRR---HVLPAQQLKQSGSLANLAVEDSDC-- 188

179 NSDTMYIADKEGGLIVHNSDDSPHRLTSNTFDYDPKFTKMTIDGESYTAODGSGWA 238  
 189 -ENTFAYADLGLVGLVSWKDEESWRVQHFFHFDPMAGNSINGIEFQWDDGLYGLA 247  
 239 LS-PMTN---NLXYSVPASTSLYYVNTTEQRTSDYQONDIIHYEGVQNI-----LDTQSSAK 290  
 248 LSKPLETGATLVPFLCSTTESVDTSLRNKTLATSPMYREFKVLGSRGNTQAGAE 307  
 291 VVS-KSGVLPFGVLGDSALGCWNEHRTLEHRTNRTVAQSDTETLQMIASMKIKKAXPHVPI 349  
 308 FLOPDGTGLVFPALPNLNEVACWRTATDFSHSSQSRIHMNDTLVFPSPDIKVD-----360  
 350 FDIYINREYILVLSNKMOKMVNDNFDDVFRIMANVNELILNTRCE 398  
 361 -----QKRLMWLSNQLPVFIYDELVYAGSINFRILTASVKEAIENTACE 403

RESULT 4  
 US-11-097-143-25233  
 ; Sequence 25233, Application US/11097143  
 ; Publication No. US20050208558A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Venter, J. Craig  
 ; APPLICANT: et al.  
 ; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
 ; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
 ; FILE REFERENCE: CL000728  
 ; CURRENT APPLICATION NUMBER: US/11/097,143  
 ; CURRENT FILING DATE: 2005-04-04  
 ; PRIOR APPLICATION NUMBER: 60/157,832  
 ; PRIOR FILING DATE: 1999-10-05  
 ; PRIOR APPLICATION NUMBER: 60/160,191  
 ; PRIOR FILING DATE: 1999-10-19  
 ; PRIOR APPLICATION NUMBER: 60/161,932  
 ; PRIOR FILING DATE: 1999-10-28  
 ; PRIOR APPLICATION NUMBER: 60/164,769  
 ; PRIOR FILING DATE: 1999-11-12  
 ; PRIOR APPLICATION NUMBER: 60/173,383  
 ; PRIOR FILING DATE: 1999-12-28  
 ; PRIOR APPLICATION NUMBER: 60/175,693  
 ; PRIOR FILING DATE: 2000-01-12  
 ; PRIOR APPLICATION NUMBER: 60/184,831  
 ; PRIOR FILING DATE: 2000-02-24  
 ; PRIOR APPLICATION NUMBER: 60/191,637  
 ; PRIOR FILING DATE: 2000-03-23  
 ; NUMBER OF SEQ ID NOS: 43008  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 25233  
 ; LENGTH: 426  
 ; TYPE: PRT  
 ; ORGANISM: DROSOPHILA  
 US-11-097-143-25233

Query Match 19.4%; Score 422; DB 6; Length 426;  
 Best Local Similarity 26.8%; Pred. No. 9.3e-29;  
 Matches 112; Conservative 83; Mismatches 171; Indels 52; Gaps 13;  
 10 KSLPILHEWKFDDYDFGSDERQDALLSGEYDYKNYPSDIDOWH-----DKIFVTMLR 63  
 30 RDTPIVFEWKNLQGFPEQERQDLNNGRNPSPIDIDIVYVPPNGGPPRHFVTSR 89  
 64 Y-NGVPSLNVISKKVGGGLLQPY--DWSFAKYDDCSGIVSASKLAIDKCDRLWYLD 120  
 90 FQGGVPSLGVTVNVQRENGEIQAYPSYQHSSHGANGCDGLTSYVRVHIDACGQWYLD 149  
 121 SGLVNNTPMCSKLLTFLDITTSOLLKQVEIPHVDVANNATTGKRLLSSLAQSDLCNTN- 179  
 150 SGEIEFVQ-HCAQVQWVFLATDQLIHYRLPE-----TSYKAVSRFVNIFADIRDP 202  
 180 -----SDTWYIADEKGLIVHNSDDSPHRLTSNTFDY-DPKFTKMTIDGESYTAODG 233

203 PSQCCKDFAYLADPTSKAIVVYDVVQSSWRI-ENKFTYDPAKFGTHTVAGSFELLGD 261  
 234 ISGMALSP---MTNNLYSPVASTSLXYVNTTEQRTSDYQONDIIHYE-GVQVILD---- 284  
 262 PLAMATPIGLGLURHLIFHALSN-----ELELAIPDLILNATNWKQKGLSSLSSEFTV 315  
 285 -----TQSSAKVSKSGVLFFGLVGDSSALGCWNEHRTLEHRTNRTVAQSDTETLQMIASMK 339  
 316 LGRGIQASHAISRQGLFCGFLPIGFIWDIRRPYNRENKLLAINPATLQFVSGMK 375  
 340 IKKAXPHVIFDIYINREYILVLSNKMOKMVNDNFDDVFRIMANVNELILNTRC 397  
 376 IVERPAD-----GREELWLLSDRLQKIFAGTIDYREINRYVMRCDDVLLQGRCC 425

RESULT 5  
 US-11-097-143-20991  
 ; Sequence 20991, Application US/11097143  
 ; Publication No. US20050208558A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Venter, J. Craig  
 ; APPLICANT: et al.  
 ; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
 ; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
 ; FILE REFERENCE: CL000728  
 ; CURRENT APPLICATION NUMBER: US/11/097,143  
 ; CURRENT FILING DATE: 2005-04-04  
 ; PRIOR APPLICATION NUMBER: 60/157,832  
 ; PRIOR FILING DATE: 1999-10-05  
 ; PRIOR APPLICATION NUMBER: 60/160,191  
 ; PRIOR FILING DATE: 1999-10-19  
 ; PRIOR APPLICATION NUMBER: 60/161,932  
 ; PRIOR FILING DATE: 1999-10-28  
 ; PRIOR APPLICATION NUMBER: 60/164,769  
 ; PRIOR FILING DATE: 1999-11-12  
 ; PRIOR APPLICATION NUMBER: 60/173,383  
 ; PRIOR FILING DATE: 1999-12-28  
 ; PRIOR APPLICATION NUMBER: 60/175,693  
 ; PRIOR FILING DATE: 2000-01-12  
 ; PRIOR APPLICATION NUMBER: 60/184,831  
 ; PRIOR FILING DATE: 2000-02-24  
 ; PRIOR APPLICATION NUMBER: 60/191,637  
 ; PRIOR FILING DATE: 2000-03-23  
 ; NUMBER OF SEQ ID NOS: 43008  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 20991  
 ; LENGTH: 412  
 ; TYPE: PRT  
 ; ORGANISM: DROSOPHILA  
 US-11-097-143-20991

Query Match 19.1%; Score 415; DB 6; Length 412;  
 Best Local Similarity 29.7%; Pred. No. 3.7e-28;  
 Matches 120; Conservative 71; Mismatches 161; Indels 52; Gaps 14;  
 22 DYDFGSDERQDALLSGEYDYKNYPSDIDOWH-----DKIFVTMLRY-NGVPSLNV 74  
 32 ELEFPSPQERQALRDGLDYGSPVIDVDVYVYKHGATPSIFVTIPFAKGVPSLAYV 91  
 75 SKKVGDDGGLLQPY--DWSFAKYDDCSGIVSASKLAIDKCDRLWYLDVSLVNNTPMCS 132  
 92 TNEVRPNTGLLOAYPSYVSEWHKSHGANGCDGLTSYVRVHIDACGQWYLDGIDFIQ-HCP 150  
 133 PKLATFDITTSOLLKQVEIPHVDVANNATTGKRLLSSLAQSDLCNTNSDTWYIADEKGE 192  
 151 POLVAIDLESCKVAHQVQMPKRL---YKEGVSRFVTPTVE-LDPHNCVDFVYVYMAISGD 206  
 193 GLIYVHNSDDSPHRLTSNTFDY-DPKFTKMTIDGESYTAQDGISGMALSPM-----TNLY 247  
 207 GIVYDVAAQSSWRI-ENKFTYDPAKFGTHTVAGSFELLGD 265  
 248 YSPVAS-----TSLYVNTTEQRTSDYQONDIIHYEGVQVILD-----TQSSAKVVS 293

Db 266 FHSLSSEWQMAIPLDVVN-----NGSNWRLND-----VSAALDQFOLLKKGKSGCQVAAAMS 316  
QY 294 KSGVLFGLGDSALGCWNEHRTLERINIRTVAGSDTQLQMIASMKIEXAPHVIPIDRY 353  
Db 317 ESGFLICGLQPSALLAWNIRTVGSHQNLVWLVEDEQLQFASGLKIVR-----NH 367  
QY 354 INREYILVLSNKKQKVVNNDFNFDVNNFRIMNANVNELIINTRC 397  
Db 368 EGKEELWLSNRLQKAFAGGLDYKEINFRIQKCGVQELLSGRPC 411

## RESULT 6

US-11-097-143-20973  
; Sequence 20973, Application US/11097143  
; Publication No. US20050208558A1  
; GENERAL INFORMATION:  
; APPLICANT: Venter, J. Craig  
; APPLICANT: et al.  
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
; TITLE OF INVENTION: DROSOPHILA GENES.  
; FILE REFERENCE: CL000728  
; CURRENT APPLICATION NUMBER: US/11/097,143  
; CURRENT FILING DATE: 2005-04-04  
; PRIOR APPLICATION NUMBER: 60/157,832  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: 60/160,191  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR APPLICATION NUMBER: 60/161,932  
; PRIOR FILING DATE: 1999-10-28  
; PRIOR APPLICATION NUMBER: 60/164,769  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR APPLICATION NUMBER: 60/173,383  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: 60/175,693  
; PRIOR FILING DATE: 2000-01-12  
; PRIOR APPLICATION NUMBER: 60/184,831  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/191,637  
; PRIOR FILING DATE: 2000-03-23  
; NUMBER OF SEQ ID NOS: 43008  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 20973  
; LENGTH: 415  
; TYPE: PRT  
; ORGANISM: DROSOPHILA  
US-11-097-143-20973

Query Match 18.8%; Score 411.5; DB 6; Length 415;  
Best Local Similarity 28.0%; Pred. No. 7.7e-28;  
Matches 119; Conservative 74; Mismatches 147; Indels 85; Gaps 15;  
QY 5 GESLNSKLPT-----LHEWKFDDYDFGSDRRQDAIISGEYDYKNNYVPSIDQWH----- 54  
Db 19 GVGNGRPRYKTVETLTQWGLEFGFTAQDRENAQAAGNLVPENGTPIDVQPYMWANGQ 78  
QY 55 DKIFVTMLRY-NGVPSLSNIVSKVGGGGLLPQPD--WSFAKYDDCSGIVSASKLAID 111  
Db 79 IRLFTTIPRTVGTGIPYLTATVSATQNGEFLLPQYPNYSWHNANGEDCCDITSAFVAIT 138  
QY 112 KCDRLWYLDGLVNNTQPMCSKLLTPTDTSOLLKQVEIPHVDVANNATTGKRL-----S 167  
Db 139 ECNQMWVIDSGVIGTTQ-LCPQQLQPALATDRLHRFRFNDTYIPS-----GSLFITPN 193  
QY 168 SLAVQSLDCNTNSDTMYIADKGEGLIVVHNSDDSHRLTSNTFDY-DPKFTKMTIDGE 226  
Db 194 VLVDPPPRGTCRMTIYVADTSW-----AENRFMYDPDPYKGHTIAGE 238  
QY 227 SYTAQGISGALSPMTNNLYSPVASTSLYYVN-----TEQFRITSY 269  
Db 239 SFYLMGDM--FALNNDKRNLYFPFLASASEYSVPLSALNRQONWANGPEALPEEFRLGR 296

QY 270 QQNDIHYEGVQNTLDQTSSAKVYKSGVLFPGVLGVSALGCWNEHRTLERINIRTVAGSD 329  
Db 297 RRS-----ECAASAIIDGNVNYCVTFNPVKLF-----VWVNSPYNSRNFGLNPAKS 343  
QY 330 ETQLQMIASMKIEXAPHVIPIDRYINR---EYILVLSNKKQKVVNNDFNFDVNNFRIMNA 386  
Db 344 DDLQFVSGMKVLR-----NREGQEELWLSNRYQKIAAGTILNSKEVNFRLRR 391  
QY 387 NVNEL 391  
Db 392 KLDDV 396

## RESULT 7

US-11-097-143-28038  
; Sequence 28038, Application US/11097143  
; Publication No. US20050208558A1  
; GENERAL INFORMATION:  
; APPLICANT: Venter, J. Craig  
; APPLICANT: et al.  
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
; TITLE OF INVENTION: DROSOPHILA GENES.  
; FILE REFERENCE: CL000728  
; CURRENT APPLICATION NUMBER: US/11/097,143  
; CURRENT FILING DATE: 2005-04-04  
; PRIOR APPLICATION NUMBER: 60/157,832  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: 60/160,191  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR APPLICATION NUMBER: 60/161,932  
; PRIOR FILING DATE: 1999-10-28  
; PRIOR APPLICATION NUMBER: 60/164,769  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR APPLICATION NUMBER: 60/173,383  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: 60/175,693  
; PRIOR FILING DATE: 2000-01-12  
; PRIOR APPLICATION NUMBER: 60/184,831  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/191,637  
; PRIOR FILING DATE: 2000-03-23  
; NUMBER OF SEQ ID NOS: 43008  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 28038  
; LENGTH: 429  
; TYPE: PRT  
; ORGANISM: DROSOPHILA  
US-11-097-143-28038

Query Match 16.8%; Score 366; DB 6; Length 429;  
Best Local Similarity 27.6%; Pred. No. 9.5e-24;  
Matches 105; Conservative 68; Mismatches 169; Indels 38; Gaps 12;  
QY 37 SGFYDYKNNYVPSIDQWHDKIFVTM-LRYNGVPSLSNVI--SKKVGDDGGLLPQYDWSF 93  
Db 70 SSGFIQVNNYVPSIDQWHDKIFVTM-LRYNGVPSLSNVI--SKKVGDDGGLLPQYDWSF 129  
QY 94 AKYDDC-SGIVSASKLAIDKCDRLWYLDGLV---NNTQPMCSKLLTPTDTSOLLKQV 149  
Db 130 NQVNASQNLVSVYRVSVDVCGRLWFDVTGMLBFFPNRQIRHPSIWIVIDLANDRLKRF 189  
QY 150 EIPHDVAVNATTGKRLS-SLAVQSLDCNTNSDTMYIADKGEGLIVVHNSDDSHRLT 208  
Db 190 EIFQSI---VEIGKGLASITIDVGARRCN---DAYAYIPDLVNRRLHVYHLRSDRIWSFE 243  
QY 209 SNTFDYDPKFTKMTIDGESYTAODGISGMALSPM-----TNLYSPVASTSLYYVNT--- 261  
Db 244 HSFNFDPDLNLDNINIGQTFRWDGIFSAFSLGSKPDGSDRDFVFFHFMASNFVVSNRVL 303  
QY 262 -EQFRT--SDYQQNDIHYEGVQNTLDQTSSAKVYKSGVLFPGVLGVSALGCWNEHRTLE 318  
Db 304 QQEFNAARSDH-GDDFHLLGTRGPGSTQSTMHKYDPRGTGVIFFAEVQKSGVGCWKTSKPF 362







```

; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT FILING DATE: 2005-04-04
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR FILING DATE: 1999-10-05
; PRIOR FILING DATE: 1999-10-19
; PRIOR FILING DATE: 1999-10-28
; PRIOR FILING DATE: 1999-11-12
; PRIOR FILING DATE: 1999-12-28
; PRIOR FILING DATE: 2000-01-12
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32673
; LENGTH: 382
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-32673

Query Match      10.7%; Score 233; DB 6; Length 382;
Best Local Similarity 24.3%; Pred. No. 6.3e-12;
Matches 83; Conservative 62; Mismatches 143; Indels 54; Gaps 12;

Qy 11 SLPLHEWKFDFDGSERRQDAILSGEYDYKNNYPSDIDQMDHDKIFVTMLRY-NGVPS 69
Db 36 SQPI--QMTGGQFEFFPCASTKSLFKSGKFPKPKVIATRAQLIGDTIYALPRYKGVPA 93
Qy 70 SLNVISKVGGGGLLPYDPDWSKAYDDCGSIVSASKLAIDKCDRLWLDGLVNTQP 129
Db 94 TLVTSIKPGCTCTTFKPYPCWDLQEEGNCALQSVVDLVVDQNEVLWLDGIVNTLET 153
Qy 130 ---MCSKLLTFDITTSOLLKQVEIPHDVAVNATGKRLSSLAQVSLDCNTNSDTMYI 186
Db 154 PVRKCPKPVVAMSVKTKGLTKVLSLE-----GLTSSNSRLQYLIV--DYAPDGGCFYV 205
Qy 187 ADEKGEGLIVYHNSDDSFHR-----LTSNTFDYDPKFTKMTIDGESYTAQDGISGMALS 240
Db 206 SDAANRAIIVYNLQADRGFRVVLPAKVTAGCRSDVLY-----IALIRRD 250
Qy 241 PMTNLYSPVASTLSLYVNTTEQRTSDYQNDIHVEGVQN--ILD--TQSSAKV---S 293
Db 251 CGSTELYFTVLTSLNKLFSLKSEYLR-----GVADGRILDGKPSRMVIGTD 299
Qy 294 KSGVLFCLGDSALGCWNEHRTLERHNIRTVACSDTELQMI 335
Db 300 NSGAIFPRNEGDAVYRWDNTSTFVEANFKPVYRS-QTCQLV 340

RESULT 14
US-10-481-180-17
; Sequence 17, Application US/10481180
; Publication No. US20040171821A1
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS
; APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND
; APPLICANT: HUMAN SERVICES
; APPLICANT: Valenzuela, Jesus G.
; APPLICANT: Belkaid, Yasmine
; APPLICANT: Kamhawi, Shaden
; APPLICANT: Sacks, David
```

```

; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT FILING DATE: 2005-04-04
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR FILING DATE: 1999-10-05
; PRIOR FILING DATE: 1999-10-19
; PRIOR FILING DATE: 1999-10-28
; PRIOR FILING DATE: 1999-11-12
; PRIOR FILING DATE: 1999-12-28
; PRIOR FILING DATE: 2000-01-12
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 699
; LENGTH: 393
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-699

Query Match      11.8%; Score 258; DB 6; Length 393;
Best Local Similarity 25.2%; Pred. No. 3.8e-14;
Matches 79; Conservative 58; Mismatches 124; Indels 52; Gaps 12;

Qy 35 ILSEYDYKNNYPSDIDQMDHDKIFVTMLRY-NGVPSLNVISKVGGGLLPYDWSF 93
Db 68 VQSGYVPRNVIVTRAQLQDSAFVALPRYKQGVPTFLGKYNLKKGBCLTKIAPYPCWAI 127
Qy 94 AKYDDCSGIVSASKLAIDKCDRLWLDGLVNN--TOPM--CSPKLLTFDITTSOLLKQVE 150
Db 128 QEEGNCQALQSVVDIAVDQNGLLWALDVGIVNTLEQPIRRCSPKIVAINTANHKVKSID 187
Qy 151 IPHDVAVNATGKRLSSLAQVSLDCNTNSDTMYIADKGEGLIVYHNSDDSFHRLTSN 210
Db 188 LS-----DLVTSESLQFIVV---DYSKDKPFVYVADAGARSILVYDITGNKSVRLV- 238
Qy 211 TFDYDPKFTKMTIDGESYTA---ODGISGMALSPMTNLYSPVASTLSLYVNTTEQRT 266
Db 239 ----PKATAPTSD-VLYVALTSKPDG-----TSTLFFSVLSLSPRLYSIRGEYLRV 283
Qy 267 SDYQNDIHVEGVQNILDTQSAKVVSKGV-----LPFGLVGDLSALGCWNEHRTL 317
Db 284 G-----QGAGSIIDV--GPKPYGKQAVLLGADGTSUFFRYKGENDIYLWDSCTCF 332
Qy 318 ERHNIRTVACSD 330
Db 333 KAANLQEVQRGD 345

RESULT 13
US-11-097-143-32673
; Sequence 32673, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT FILING DATE: 2005-04-04
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR FILING DATE: 1999-10-05
; PRIOR FILING DATE: 1999-10-19
; PRIOR FILING DATE: 1999-10-28
; PRIOR FILING DATE: 1999-11-12
; PRIOR FILING DATE: 1999-12-28
; PRIOR FILING DATE: 2000-01-12
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 699
; LENGTH: 393
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-32673
; Sequence 32673, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
```



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2006, 19:12:24 ; Search time 23.7942 Seconds  
(without alignments)  
803.371 Million cell updates/sec

Title: US-10-525-567-4

Perfect score: 2178

Sequence: 1 NILRGESLNKSLPILHEWKF.....NTRCENPDNRTFPKISHL 413

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 235405 seqs, 46284737 residues

Total number of hits satisfying chosen parameters: 235405

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:

- 1: /SIDSS/ptodata/1/pubpaa/US08 NEW PUB.pep.1\*
- 2: /SIDSS/ptodata/1/pubpaa/US06 NEW PUB.pep.\*
- 3: /SIDSS/ptodata/1/pubpaa/US07 NEW PUB.pep.\*
- 4: /SIDSS/ptodata/1/pubpaa/US08 NEW PUB.pep.\*
- 5: /SIDSS/ptodata/1/pubpaa/PCT NEW PUB.pep.\*
- 6: /SIDSS/ptodata/1/pubpaa/US09 NEW PUB.pep.\*
- 7: /SIDSS/ptodata/1/pubpaa/US09 NEW PUB.pep.1\*
- 8: /SIDSS/ptodata/1/pubpaa/US10 NEW PUB.pep.\*
- 9: /SIDSS/ptodata/1/pubpaa/US10 NEW PUB.pep.1\*
- 10: /SIDSS/ptodata/1/pubpaa/US11 NEW PUB.pep.\*
- 11: /SIDSS/ptodata/1/pubpaa/US11 NEW PUB.pep.1\*
- 12: /SIDSS/ptodata/1/pubpaa/US60 NEW PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	276.5	12.7	397	9	US-10-533-811-29
2	260	11.9	388	9	Sequence 29, Appl
3	256	11.8	388	9	Sequence 51, Appl
4	240.5	11.0	393	9	Sequence 5, Appl
5	240.5	11.0	412	9	Sequence 7, Appl
6	214	9.8	393	9	Sequence 23, Appl
7	184	8.4	399	9	Sequence 49, Appl
8	113.5	5.2	1530	11	Sequence 63, Appl
9	108.5	5.0	829	9	Sequence 34, Appl
10	100.5	4.6	2314	11	Sequence 26, Appl
11	100.5	4.6	2353	11	Sequence 2, Appl
12	99.5	4.6	2376	11	Sequence 6, Appl
13	99.5	4.6	2715	11	Sequence 4, Appl
14	97.5	4.5	371	9	Sequence 2, Appl
15	97.5	4.5	371	9	Sequence 6, Appl
16	97.5	4.5	371	9	Sequence 4, Appl
17	97.5	4.5	371	9	Sequence 2, Appl
18	97.5	4.5	371	9	Sequence 146, Appl
19	97.5	4.5	371	9	Sequence 146, Appl
20	97.5	4.5	371	9	Sequence 146, Appl
21	97.5	4.5	371	9	Sequence 144, Appl
22	97.5	4.5	371	9	Sequence 144, Appl

22	97.5	4.5	371	9	US-10-233-134-144	Sequence 144, Appl
23	97	4.5	490	11	US-11-096-568A-30481	Sequence 30481, A
24	97	4.5	558	11	US-11-096-568A-30480	Sequence 30480, A
25	97	4.5	690	11	US-11-096-568A-30479	Sequence 30479, A
26	96.5	4.4	434	11	US-11-096-568A-24278	Sequence 24278, A
27	96.5	4.4	451	11	US-11-096-568A-24277	Sequence 24277, A
28	96.5	4.4	470	11	US-11-096-568A-24276	Sequence 24276, A
29	95.5	4.4	505	11	US-11-087-099-7776	Sequence 7776, Ap
30	95.5	4.4	2613	9	US-10-455-772-530	Sequence 530, App
31	95.5	4.4	2628	9	US-10-455-772-502	Sequence 502, App
32	95.5	4.4	2715	11	US-11-113-424-51	Sequence 51, Appl
33	95.5	4.4	2721	9	US-10-455-772-522	Sequence 10, Appl
34	95.5	4.4	2721	11	US-11-096-051-10	Sequence 486, App
35	95.5	4.4	2725	9	US-10-455-772-486	Sequence 536, App
36	95.5	4.4	2725	9	US-10-455-772-526	Sequence 544, App
37	95.5	4.4	2725	9	US-10-455-772-544	Sequence 546, App
38	95.5	4.4	2725	9	US-10-455-772-546	Sequence 548, App
39	95.5	4.4	2725	9	US-10-455-772-548	Sequence 550, App
40	95.5	4.4	2725	9	US-10-455-772-550	Sequence 552, App
41	95.5	4.4	2725	9	US-10-455-772-552	Sequence 8, Appl1
42	95.5	4.4	2725	11	US-11-096-051-8	Sequence 8475, Ap
43	95	4.4	483	11	US-11-079-463-8475	Sequence 1142, Ap
44	94.5	4.3	4913	9	US-10-453-372-1142	Sequence 1132, Ap
45	94.5	4.3	4961	9	US-10-453-372-1132	

ALIGNMENTS

RESULT 1

US-10-533-811-29  
; Sequence 29, Application US/10533811  
; Publication No. US20060051364A1  
; GENERAL INFORMATION:  
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS  
; APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND  
; APPLICANT: HUMAN SERVICES  
; APPLICANT: Valenzuela, Jesus G.  
; APPLICANT: Ribeiro, Jose M.C.  
; APPLICANT: Barral, Aldina  
; APPLICANT: Netto, Manoel  
; APPLICANT: Brodskyn, Claudia  
; APPLICANT: Gomes, Regis  
; TITLE OF INVENTION: LU. LONGIPALPIS POLYPEPTIDES AND METHODS OF USE  
; FILE REFERENCE: 4239-67028-08  
; CURRENT APPLICATION NUMBER: US/10/533,811  
; CURRENT FILING DATE: 2005-04-29  
; PRIOR FILING DATE: 2002-10-29  
; PRIOR APPLICATION NUMBER: PCT/US2003/034453  
; PRIOR FILING DATE: 2003-10-29  
; NUMBER OF SEQ ID NOS: 73  
; SOFTWARE: Patent in version 3.3  
; SEQ ID NO 29  
; LENGTH: 397  
; TYPE: PRT  
; ORGANISM: Lutzomyia longipalpis  
; US-10-533-811-29

Query Match 12.7%; Score 276.5; DB 9; Length 397;

Best Local Similarity 24.9%; Pred. No. 1.2e-15;

Matches 107; Conservative 68; Mismatches 165; Indels 89; Gaps 20;

QY	20	FFDYDFGSDER-----RQDAILSGEYDYKKNYPSD-----IDQWHDKIFVTM 61
Db	5	FFLYTFGLVQITFGVEIKQGFKNKILY-EGDTSENFNPNNILTAIFYAPDESQKLFIV 63
QY	62	LYRNGVPSLNVISK-----KVGDDGGLLPQYPDMSPAKYDDCSGIVSASKLAIDKC 113
Db	64	PR--KYPETMTLAEVDETEKNSFSGDTSPL-----CKFSGHGTGKLTSTVYQVIDEC 116
QY	114	DLRWLDLSGLVNNT-----QPMCSPKLLTFLDTLTSQLLKQVEIPH-----DVAVNAT 160

Db 117 HRLVWVGVSGVNSDGTGEPHNPTLVAYD-----LKEANYPEVIRVTFPDNSIEKP 170  
Qy 161 TKGRLSSALVQSLDCNTSNSTWYVIADEKGEGLIVY-HNSDDSFHRLTSNTFYDPKFT 219  
Db 171 TFLGGAADVVPKDECC--SETFVYINFLNALIVVDHKKDSW-TVQDSTFGPKK-S 225  
Qy 220 KMTIDGESYTAQDQISGMALSPWN----NLYSPVASTSLYYVNTQFRTSDYQNDI 275  
Db 226 KFDHGGQYQYEAGIFGTLGERDNEGRQAYYLVASSTKLHSINTKELQKGSKNV-AN 284  
Qy 276 YEGVQNLDTQSSAKVVS-----KSGVLFGVGLGDSALGCWNEHRTLRHNIRTVAAQSD 330  
Db 285 YLG-----DRGSTAIGLVYDPKTKTIFVFNESKRVSCWNTQETLKKIDVINYENAD 339  
Qy 331 TLQMIASMKIEXPHVPIFDYRINREYILVLSNMOKMVND-FNFDVNVFRIMNANVN 389  
Db 340 -FSFGTIDSID-----SQDNLWFLANGLPLENSDKFVTKRYQIFKYNIQ 385  
Qy 390 ELILNTRCE 398  
Db 386 EAIAGTKCE 394

## RESULT 2

US-10-527-500-51  
; Sequence 51, Application US/10527500  
; Publication No. US20060004186A1  
; GENERAL INFORMATION:  
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS  
; APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND  
; APPLICANT: HUMAN SERVICES  
; APPLICANT: Valenzuela, Jesus G.  
; APPLICANT: Ribeiro, Jose M.C.  
; APPLICANT: Kamhawi, Shaden  
; APPLICANT: Belkaid, Yasmine  
; APPLICANT: Fischer, Laurent Bernard  
; APPLICANT: Audonnet, Jean-Cristophe  
; APPLICANT: Milward, Francis William  
; TITLE OF INVENTION: P. ARIASI POLYPEPTIDES AND P. PERNICIOSUS POLYPEPTIDES AND  
; TITLE OF INVENTION: METHODS OF USE  
; FILE REFERENCE: 4239-66903-02  
; CURRENT APPLICATION NUMBER: US/10/527,500  
; CURRENT FILING DATE: 2005-03-11  
; PRIOR APPLICATION NUMBER: PCT/US2003/029833  
; PRIOR FILING DATE: 2003-09-18  
; PRIOR APPLICATION NUMBER: US 60/425,852  
; PRIOR FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: US 60/412,327  
; PRIOR FILING DATE: 2002-09-19  
; NUMBER OF SEQ ID NOS: 87  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 51  
; LENGTH: 388  
; TYPE: PRT  
; ORGANISM: Phlebotomus perniciosus  
US-10-527-500-51

Query Match 11.9%; Score 260; DB 9; Length 388;  
Best Local Similarity 23.5%; Pred. No. 2.9e-14;  
Matches 96; Conservative 67; Mismatches 172; Indels 74; Gaps 14;  
Qy 16 HEWKFFDYDFGSDERRQDAILSGEYDYKNN-----YPSDIDQWHDKIFVTMLRYNGVPSSL 71  
Db 25 YAWKNISF-----EGIDPASYSVKNSIVTGFADADS--KKIFITIPRLNPVITL 73  
Qy 72 NVI-SKKVGDGGLQPPYDPSFAKYDDCSGIVSASKLAIDKCDRLWVLSGLV-----N 125  
Db 74 TELDTTHPEGSPLSKFP-----GSDKLSIVYQVVIDECRLWIVDAGQVEYKGD 125  
Qy 126 NTQWCSKPLLTFLTLTSQL--LKQVEIPHVDVAVNATTGKRLSSLAQSLDCNTNSDTM 183  
Db 126 QKIPKNAIATYDLTKDNTPEIDRYELPNNVAGNPLGFGGFAVDVTPNPKGC---GKTF 182

Qy 184 VYIADKSGEGLIVHNSDDSFHRLTSNTFYDPKFTMTIDGESYTAQDQISGMALSPM- 242  
Db 183 VYITNFEDNTLIVYDQEKDWSKISHDSFKPEHE-SILTHNGAQHILKLGIFGITLGLD 241  
Qy 243 ---TNLNYSPVASTSLYYVNTQFRTSDYQ-----QNDIHVEGVQNLDTQSSAKV 292  
Db 242 EGNRQAYLGGSTKLFRVNTKOLKKAGQIBFTPLGRGSHSEALALAYD----- 293  
Qy 293 SKSGVLFFGLVGLGDSALGCWNEHRTLRHNIRTVAAQSDETLQMIASMKIEXPHVPIFDR 352  
Db 294 PKTKVIFFEYNSKRISQWNTQKSLNPDNDIVYHSPDF-----IFGT 336  
Qy 353 YINRE---YLVLSNMOKMVND-FNFDVNVFRIMNANVELILNTRCE 398  
Db 337 DISMDSKLPWFSGHPPIENVOLTFDPKPHFLRLISMDTKKSIHGKCE 385  
RESULT 3  
US-10-527-500-5  
; Sequence 5, Application US/10527500  
; Publication No. US20060004186A1  
; GENERAL INFORMATION:  
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS  
; APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND  
; APPLICANT: HUMAN SERVICES  
; APPLICANT: Valenzuela, Jesus G.  
; APPLICANT: Ribeiro, Jose M.C.  
; APPLICANT: Kamhawi, Shaden  
; APPLICANT: Belkaid, Yasmine  
; APPLICANT: Fischer, Laurent Bernard  
; APPLICANT: Audonnet, Jean-Cristophe  
; APPLICANT: Milward, Francis William  
; TITLE OF INVENTION: P. ARIASI POLYPEPTIDES AND P. PERNICIOSUS POLYPEPTIDES AND  
; TITLE OF INVENTION: METHODS OF USE  
; FILE REFERENCE: 4239-66903-02  
; CURRENT APPLICATION NUMBER: US/10/527,500  
; CURRENT FILING DATE: 2005-03-11  
; PRIOR APPLICATION NUMBER: PCT/US2003/029833  
; PRIOR FILING DATE: 2003-09-18  
; PRIOR APPLICATION NUMBER: US 60/425,852  
; PRIOR FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: US 60/412,327  
; PRIOR FILING DATE: 2002-09-19  
; NUMBER OF SEQ ID NOS: 87  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 5  
; LENGTH: 388  
; TYPE: PRT  
; ORGANISM: Phlebotomus ariasi  
US-10-527-500-5

Query Match 11.8%; Score 256; DB 9; Length 388;  
Best Local Similarity 24.0%; Pred. No. 6.3e-14;  
Matches 98; Conservative 63; Mismatches 175; Indels 72; Gaps 15;  
Qy 16 HEWKFFDYDFGSDERRQDAILSGEYDYKNNYPSDI--DQWHDKIFVTMLRYNGVPSSLNV 73  
Db 25 YAWKNIYE-GIDQ-----GSYNIENSIPTAFADAAASKKIFITIPRINQVPTL 75  
Qy 74 I-SKKVGDGGLQPPYDPSFAKYDDCSGIVSASKLAIDKCDRLWVLSGLV-----NNT 127  
Db 76 FDSIKYPGSGPLSKFP-----GSDNIISVYQVVIDECRLWIVDAGQVEYKGD 127  
Qy 128 QPMCSKPLLTFLTLTSQL--LKQVEIPHVDVAVNATTGKRLSSLAQSLDCNTNSDTM 185  
Db 128 YPKNPAITAYDLTKDNTPEIDRYELPINIAGNPLGFGGFTVDVTPNPKGC---GKTF 184  
Qy 186 IADEKGEGLIVHNSDDSFHRLTSNTFYDPKFTMTI-DGESYTAQDQISGMAL-----S 240  
Db 185 ITNFEDNTLIVYDQEKDWSKISHGSGF--KPEHESILHNGVDHILKLGIFGITLGRDS 242  
Qy 241 PMTNLNYSPVASTSLYYVNTQFRTSDYQNDI-----HYEGVQNLDTQSSAKV 293

Db 243 EGNRPAYLGGSTKLFVNTKALKKKEGEIEPTITLGDGRPHSEAIALAYD-----P 294

Qy 294 KSGVLFGLVGDLSALGWNHRTLRNIRTVASQDETQMIASMKIKEAXPHVIPFDY 353

Db 295 KTKVIFTEYNSKISKISWNKIKPLIHDMKDIYASPEF-----IFGTD 337

Qy 354 I---NREYILVLSNMQOMVNDNFDDVNFPRIMANVNELLNTRCE 398

Db 338 ISVDSSEKLMFFSNGHPPIENLQSSDKPHHLISVDEKAIKRTKCE 385

RESULT 4

US-10-527-500-7

Sequence 7, Application US/10527500

Publication No. US20060004186A1

GENERAL INFORMATION:

APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS

APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND

APPLICANT: HUMAN SERVICES

APPLICANT: Valenzuela, Jesus G.

APPLICANT: Ribeiro, Jose M.C.

APPLICANT: Belkaid, Yasmine

APPLICANT: Fischer, Laurent Bernard

APPLICANT: Audonnet, Jean-Cristophe

APPLICANT: Milward, Francis William

TITLE OF INVENTION: P. ARIASI POLYPEPTIDES AND P. PERNICIOSUS POLYPEPTIDES AND

TITLE OF INVENTION: METHODS OF USE

FILE REFERENCE: 4239-66903-02

CURRENT APPLICATION NUMBER: US/10/527,500

CURRENT FILING DATE: 2005-03-11

PRIOR APPLICATION NUMBER: PCT/US2003/029833

PRIOR FILING DATE: 2003-09-18

PRIOR FILING DATE: 2002-11-12

PRIOR FILING DATE: 2002-11-12

PRIOR FILING DATE: 2002-09-19

NUMBER OF SEQ ID NOS: 87

SOFTWARE: PatentIn version 3.2

SEQ ID NO 7

LENGTH: 393

TYPE: PRT

ORGANISM: Phlebotomus ariasi

US-10-527-500-7

Query Match 11.0%; Score 240.5; DB 9; Length 393;

Best Local Similarity 24.2%; Pred. No. 1.3e-12;

Matches 97; Conservative 59; Mismatches 167; Indels 77; Gaps 16;

Qy 39 EYDYKNNVPSDIDQWHDKIFVTMLRYNGVPSL--NVISKVGGDGPILQP----- 87

Db 24 EYAKNNTPEGIDQASYN-----ENSIPTAFVHDALSKKIIAIPRLYPQVPITLQ 77

Qy 88 ----YDHS--PAKYDDCSGIIVASAKLAIDKCDRLWLDGLV-----NNTQPMCSKLL 136

Db 78 DTTKHPERSPLEKPEPGSKLTSVQPMLEDCRRUWIDVQGVYKGBQKPKKPAII 137

Qy 137 TFDLTSQL--LKQVEIPHVDVAVNATTGKRLSSLAQSLDNTNMTVYIADKGBGL 194

Db 138 AYDLTKDNPYEDIRVEIPINAGNIQFGFTVDVTPNKEG--GKTFIYITNEDNTL 194

Qy 195 IVYHNSDDSFHRLTNTFP--DYDPKFTKMTIDGESYTAQDGSGLM-----SPMTN-LYY 248

Db 195 IYVDQEKDQSKWIKSHGSKFPEHESNFH---NGAQYKYKAGIFGITLGDREDPEGRN 251

Qy 249 SPVASTSLYVNTQFRTSDYQNDI-----HYEGVQNIQDQSSAKVSKVSGVLPFG 301

Db 252 LGSSTKLFVEVSTEALKKGAKFPVRLGDGRGHTAIALVYD-----PKTKVIFPA 303

Qy 302 LVGDSALGWCNHEHRTLRNIRTVASQDETQMIASMKIKEAXPHVIPFDYI---NREY 358

Db 304 ESDRSQISQNTQKPLNHN-----TDVIYASSKF-----IFGTDIQIDSQ 346

Qy 359 ILVLSNMQOMVNDNFDDVNFPRIMANVNELLNTRCE 398

Db 347 LWLSNGQPIDNKLTFQKPHIRLARVDTKNSIRTRCE 386

RESULT 5

US-10-533-811-23

Sequence 23, Application US/10533811

Publication No. US20060051364A1

GENERAL INFORMATION:

APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS

APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND

APPLICANT: HUMAN SERVICES

APPLICANT: Valenzuela, Jesus G.

APPLICANT: Ribeiro, Jose M.C.

APPLICANT: Barral, Aldina

APPLICANT: Netto, Manoel

APPLICANT: Brodskyn, Claudia

APPLICANT: Gomes, Regis

TITLE OF INVENTION: LU. LONGIPALPIS POLYPEPTIDES AND METHODS OF USE

FILE REFERENCE: 4239-67028-08

CURRENT APPLICATION NUMBER: US/10/533,811

CURRENT FILING DATE: 2005-04-29

PRIOR APPLICATION NUMBER: US 60/422,303

PRIOR FILING DATE: 2002-10-29

PRIOR APPLICATION NUMBER: PCT/US2003/034453

PRIOR FILING DATE: 2003-10-29

NUMBER OF SEQ ID NOS: 73

SOFTWARE: PatentIn version 3.3

SEQ ID NO 23

LENGTH: 412

TYPE: PRT

ORGANISM: Lutzomyia longipalpis

US-10-533-811-23

Query Match 11.0%; Score 240.5; DB 9; Length 412;

Best Local Similarity 24.7%; Pred. No. 1.4e-12;

Matches 103; Conservative 69; Mismatches 170; Indels 75; Gaps 19;

Qy 33 DAILSGEYDYKNNVPSD-IDQWHDKIFVTMLRYN-GVPSLNVISKVGGDGPILQYP 89

Db 33 DGLTDDYNPKFNPTGLAVDPEGYRLFIAPRRKPKVPYTVAEALNMVMPGFP-VERAP 91

Qy 90 DWSPAKYDDCSG-----IVSASKLAIDKCDRLWLDGLVNNT-----QPMCSFKLTF 138

Db 92 --SFEKFKFNGEGKDLVNVYQPVIDDRELWLDIGKVEYTCGDADQYKPKGPTLIAY 149

Qy 139 DLTTSQL--LKQVEIPHVDV-----AVNATTGKRLSSLAQSLDNTNMTVYI 186

Db 150 DLKXDHTEPHRPEIPDDLYSSQVEFGGFAVDVNTKG-----DC---TESFYL 196

Qy 187 ADEKGEGLIVYHNSDDSFHRLTNTFDYDPKFTKMTIDGESYTAQDGSGLMSPMTN 246

Db 197 TNFDKNSLIVYDEYTKKAWKFTDKTFEADKESTFSYSGEOMKYKGLGFGALGD-RDEM 255

Qy 247 YVSPV-----ASTSLYVNTQFRTSDYQON-DIHYEGVQNIQDQSSAKVSKVSGVLF 300

Db 256 GHRPACYTAGSTKVSNTKELKTNQNLNQLH--GDRKYTDALAYDPEHKVLYF 313

Qy 301 GLVDSALGWCNHEHRTLRNIRTVASQ-----DETQMIASMKIKEAXPHVIPFDYI 354

Db 314 AESDSRQVSCWNVNMLKPDNTDVFSSARFTFGTDLVDSKGLMWIANGHPVPEQO-- 371

Qy 355 NREYILVLSNMQOMVNDNFDDVNFPRIMANVNELLNTRCENPDNDRTPFKISI 411

Db 372 -----EKIWKM-----RFVNRKIRIMKVDTERVFKYSRC-NP-NYKPKKEIEV 412

RESULT 6

US-10-527-500-49

Sequence 49, Application US/10527500

Publication No. US20060004186A1

GENERAL INFORMATION:



; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS  
; APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND  
; APPLICANT: HUMAN SERVICES  
; APPLICANT: Valenzuela, Jesus G.  
; APPLICANT: Ribeiro, Jose M.C.  
; APPLICANT: Kanhawi, Shaden  
; APPLICANT: Belkaid, Yasmine  
; APPLICANT: Fischer, Laurent Bernard  
; APPLICANT: Audonnet, Jean-Cristophe  
; APPLICANT: Milward, Francis William  
; TITLE OF INVENTION: P. ARIASI POLYPEPTIDES AND P. PERNICIOSUS POLYPEPTIDES AND  
; TITLE OF INVENTION: METHODS OF USE  
; FILE REFERENCE: 4239-66903-02  
; CURRENT APPLICATION NUMBER: US/10/527,500  
; CURRENT FILING DATE: 2005-03-11  
; PRIOR APPLICATION NUMBER: PCT/US2003/029833  
; PRIOR FILING DATE: 2003-09-18  
; PRIOR APPLICATION NUMBER: US 60/425,852  
; PRIOR FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: US 60/412,327  
; PRIOR FILING DATE: 2002-09-19  
; NUMBER OF SEQ ID NOS: 87  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 49  
; LENGTH: 393  
; TYPE: PRT  
; ORGANISM: Phlebotomus perniciosus  
US-10-527-500-49

Query Match 9.8%; Score 214; DB 9; Length 393;  
Best Local Similarity 23.3%; Pred. No. 2.4e-10;  
Matches 97; Conservative 60; Mismatches 151; Indels 108; Gaps 19;

Qy 39 EYDKNNYPSDIDQWHDKIFVTMLRYNGV-PSSLNV-----ISKVGDGGPLL 85  
Db 24 EYAWKN-----ISYEGVDPALFNIDNIITPGVHDAINKKIFIAVPRR 66

Qy 86 QP-----YPDWS--FAKYDDCSGIVSASKLAIDKCDRLWLVDGLVN----- 125  
Db 67 SPQIPFTLTDLTTKHPSRSPPLSKFPGSDKLINVYQVIDECRLWIADVGRVYKGD 126

Qy 126 NTQPMCSKLLTFLTTSQL--LKQVEIPHDAVNAITGKRSLSSLAQSLDNTN---- 179  
Db 127 QKYENQNAVLIAYDLTKENYPEIHRYPISKAGSNTIPFG---GFAVDV---TNPKEG 179

Qy 180 -SDTMVYIADEKGBGLVYHNSDDSFHRLTSNTP--DYDPKFTQWITDGSYTAQDGISG 236  
Db 180 CGKTFVITPFDNTLIVYDQEKDQSWKISHGSKPEHD---SVLSHDGQYKYRVGLFG 236

Qy 237 MAL---SPMTNN-LYYSFVASTSLYYNTE-----QPRTSYQQNDIHYEGVQNILDT 285  
Db 237 ITLGDRLPEGNRPAYIAGSSTKLFEISTKLKKGAKFDPVNLGNRGPHTAVALVYD- 295

Qy 286 QSSAKVSKGVLPFGVLGVSALCGWNEHRTLEHNRITVAQSDETLQMIASMKIEXAP 345  
Db 296 -----PKTKVIFFAESDSRQVSCWNTQKPLNHN-----TDVIFASAKF----- 334

Qy 346 HVPIFDRIYI---NREYILVLSNKKQKMWVNDNFDDVNFPRIMANVELINTRCE 398  
Db 335 ---IYSGDISVDSSEQLWFSTGHGHPPIPNLKLTFDKPHIRLMRVDTAKAIRTRCE 387

RESULT 7  
US-533-811-63  
; Sequence 63, Application US/10533811  
; Publication No. US20060051364A1  
; GENERAL INFORMATION:  
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS  
; APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND  
; APPLICANT: HUMAN SERVICES  
; APPLICANT: Valenzuela, Jesus G.  
; APPLICANT: Ribeiro, Jose M.C.  
; APPLICANT: Barrai, Aldina

; APPLICANT: Netto, Manoel  
; APPLICANT: Brodskyn, Claudia  
; APPLICANT: Gomes, Regis  
; TITLE OF INVENTION: LU. LONGIPALPIS POLYPEPTIDES AND METHODS OF USE  
; FILE REFERENCE: 4239-67028-08  
; CURRENT APPLICATION NUMBER: US/10/533,811  
; CURRENT FILING DATE: 2005-04-29  
; PRIOR APPLICATION NUMBER: US 60/422,303  
; PRIOR FILING DATE: 2002-10-29  
; PRIOR APPLICATION NUMBER: PCT/US2003/034453  
; PRIOR FILING DATE: 2003-10-29  
; NUMBER OF SEQ ID NOS: 73  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 63  
; LENGTH: 399  
; TYPE: PRT  
; ORGANISM: Latzomyia longipalpis  
US-10-533-811-63

Query Match 8.4%; Score 184; DB 9; Length 399;  
Best Local Similarity 22.6%; Pred. No. 8.5e-08;  
Matches 97; Conservative 61; Mismatches 168; Indels 104; Gaps 20;

Qy 16 HEWKFFDYDFGSDERRQDAILSGEYDYKKNYPSDI--DQWHDKIFVTMLR-YNGVPSLN 72  
Db 24 YKWKQLLY-----NNVTFGSYNPDNMI STAFAYDAEGEKLFLAVPRKLPVRYPTLA 74

Qy 73 VISKK-----VGDGGLLPQYPDWSPFAKYDDCSGIVSASKLAIDKCDRLWLVDGLVN-- 125  
Db 75 EVDTONSLGVKGKHSPLN-----KFSGHKTGKELTSIQPVIDDCCRLRLWVDIGSVEYR 129

Qy 126 ----NTQPMCSKLLTFLTTSQLKQVEIPHNV-----AVNATTGK 163  
Db 130 SRGAKDVPSHRPAIVAYD-----LKQPNVPEVRYVFPTRLVEKPTYFGGFAVDVANPK 183

Qy 164 GRUSSLAQSLDNTNSTDWVIADKGLVY--HNSDDSFHRLTSNTPFDYDPKTKMT 222  
Db 184 G-----DC---SETFVITNFRGALFIYDHKKQDSMN-VTHPTFKAB-RPTKFD 228

Qy 223 IDGESYTAQDGISGMAL-----SPMTNNLYSPVASTSLYYNTEQPRTSYQQNDIHYEG 278  
Db 229 YGKEVEYFAGIFGIIIGDRDSEGNRPAYLAGSAIKYVSNTKELKQGGKLN-ELLG 287

Qy 279 VQNILDTQSSAKVSKGVLPFGVLGVSALCGWNEHRTLEHNRITVAQSDETLQMIASM 338  
Db 288 NRKYNDAIALAYDPKTKVIFFAEANTQVSCWN-----TQKMPLRM 329

Qy 339 KIKEAXPHVPIFDRIYI-----NREYILVLSNKKQKMWVND-FNFDVNFPRIMN-ANV 388  
Db 330 KNTDV---VYTSRFRVFGTDISVDSKGLWFMWSNGPPIRKSEKFKYDFRYRLMRIMDT 386

Qy 389 NELILNTRCE 398  
Db 387 OEAIAGTACD 396

RESULT 8  
US-11-045-004-34  
; Sequence 34, Application US/11045004  
; Publication No. US20060078901A1  
; GENERAL INFORMATION:  
; APPLICANT: BUCHRIESER, CARMEN  
; APPLICANT: FRANGEUL, LIONEL  
; APPLICANT: COUVE, ELISABETH  
; APPLICANT: RUSNIOK, CHRISTOPHE  
; APPLICANT: FSIHI, HAFIDA  
; APPLICANT: DEHOUX, PIERRE  
; APPLICANT: DUSURGET, OLIVIER  
; APPLICANT: CHETOUANI, FARID  
; APPLICANT: NEDJARI, HAFED  
; APPLICANT: GLASER, PHILIPPE  
; APPLICANT: KUNST, FRANK  
; APPLICANT: COSSART, PASCALE



APPLICANT: DANIELS, JUSTIN  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KREFT, JURGEN  
APPLICANT: KUHN, MICHAEL  
APPLICANT: NG, EVA  
APPLICANT: VAZQUEZ-BOLAND, ANTONIO  
APPLICANT: DOMINGUEZ-BERNAL, GUSTAVO  
APPLICANT: GARRIDO-GARCIA, PATRICIA  
APPLICANT: TIERREZ-MARTINEZ, ALBERTO  
APPLICANT: AMEND, ALEXANDRA  
APPLICANT: CHAKRABORTY, TRINAD  
APPLICANT: DOMANN, EUGEN  
APPLICANT: HAIN, THORSTEN  
APPLICANT: BERCHE, PATRICK  
APPLICANT: CHARBIT, ALAIN  
APPLICANT: DURANT, LIONEL  
APPLICANT: PEREZ-DIAZ, JOSE-CLAUDIO  
APPLICANT: BAQUERO, FERNANDO  
APPLICANT: GARCIA DEL PORTILLO, FRANCISCO  
APPLICANT: GOMEZ-LOPEZ, NURIA  
APPLICANT: MADUENIO, ENCARN  
APPLICANT: PABLOS, BETRIZ DE  
APPLICANT: WEHLAND, JURGEN  
APPLICANT: KARST, UWE  
APPLICANT: ENTIAN, KARL-DIETER  
APPLICANT: HAUF, JORG  
APPLICANT: ROSE, MATTHIAS  
APPLICANT: VOSS, HAMUT  
TITLE OF INVENTION: LISTERIA MONOCYTOGENES GENOME, POLYPEPTIDES AND USES  
FILE REFERENCE: 05394.0018-02  
CURRENT APPLICATION NUMBER: US/11/045.004  
CURRENT FILING DATE: 2005-01-28  
PRIOR APPLICATION NUMBER: 10/637,657  
PRIOR FILING DATE: 2003-08-11  
PRIOR APPLICATION NUMBER: 10/257,023  
PRIOR FILING DATE: 2002-10-08  
PRIOR APPLICATION NUMBER: PCT/FR01/01118  
PRIOR FILING DATE: 2001-04-11  
PRIOR APPLICATION NUMBER: FR 00/04,629  
PRIOR FILING DATE: 2000-04-11  
NUMBER OF SEQ ID NOS: 2854  
SOFTWARE: Patentin version 3.3  
SEQ ID NO 34  
LENGTH: 1530  
TYPE: PRT  
ORGANISM: Listeria monocytogenes  
US-11-045-004-34

Query Match 5.2%; Score 113.5; DB 11; Length 1530;  
Best Local Similarity 20.8%; Pred. No. 0.55;  
Matches 106; Conservative 79; Mismatches 182; Indels 143; Gaps 32;

Qy 3 LRGESLNKSLPLHEW-KFFDYDFGSDERRQDAI-----LSGEYDYKNN----- 45  
Db 522 LNGDS-----SWEKGFQINFNRDITGQHVITYKTVQNPSTHTSGDNEYTNATIKT 572

Qy 46 ---YPSDID-QWHDKIFVTMLRYNGVSPSSLVNISKVGDGGLLPQVPDWSFAKYDDCSG 101  
Db 573 DTAERASDSTKWIDKVIDADGYKNG-----VFNYKTGE-----IEWKL1-FNDSSK 617

Qy 102 IVSASKLAID---KCDRLWLDLGLVWNTQPMCSPL-----LTFDLT-----TSOL 145  
Db 618 LI--SKPTIEDSLNSGQTFIQDSIBHKIDLSATPQVGEIAPPENYDVTFTKNGNGEQM 675

Qy 146 LKQVEIP--HDVAVNATT-----GKRLSSLAVQLDNCNTNSDTWVY 185  
Db 676 LITFKKLIHPVEVYTKTPGVCITKPLYNKRAVISDGEVLADYEAVIDDNANK-----Y 731

Qy 106 I---ADEKGEGL--IVVHNSDDSFHRLTSNFTDYPDKFTKMTIDGES---YTAQDQISG- 236  
Db 732 VNKSGEVDNDIWEIYANQSGS---TVSNATVDTLTGQKLDFTSSIKVYKSQTSVTGK 788

Qy 237 -----MALSPMTNNLYSPVASTSLYYVNTPEFTSDYQQNDIHYE-GVQNILDTQSSA 289

Db 789 MLOESNMPISGCEYDLKTGVDSESNLEYPQV-KFRNEINQSYVIKYQTAITLTSOTETTA 847  
Qy 290 KVVSKSVLFFGLVGDSALGCWNEHRTLERRHIRT-VAQSD-----ETLQMIASMKIKEA 343  
Db 848 QI--GNSVTF---TGUNI-----TKGETEKTKEIVKIITGDTGTGCTGKIILN-KVDKA 897  
Qy 344 XPHVPI-----FDRYINREYILVLSNKKMQQVNV-NDFNPDVNVFRINNVANNLILNTRCE 398  
Db 898 DSIPLLEGATFDLYANDEKVDQTDTTKNGVIGFDDLVYGDYTLKVSAPGEGYTLPTASTE 957  
Qy 399 NPD-----NDRTPFKIS---IHL 413  
Db 958 NIQVKLEQDEKVVQVNMKEKPIKETGEVHL 987

RESULT 9  
US-10-909-769-26  
; Sequence 26, Application US/10909769  
; Publication No. US20060024331A1  
; GENERAL INFORMATION:  
; APPLICANT: Fernandez-Salas, Ester  
; APPLICANT: Steward, Lance E.  
; APPLICANT: Lin, Wei-Jen  
; APPLICANT: Aoki, Kei Roger  
; APPLICANT: Sachs, George  
; TITLE OF INVENTION: Toxin Compounds with Enhanced Membrane Translocation Characteristics  
; FILE REFERENCE: ALL80010-100 (R012003-146)  
; CURRENT APPLICATION NUMBER: US/10/909,769  
; CURRENT FILING DATE: 2004-08-02  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: Patentin version 3.3  
; SEQ ID NO 26  
; LENGTH: 829  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Amino acid sequence of HC  
US-10-909-769-26

Query Match 5.0%; Score 108.5; DB 9; Length 829;  
Best Local Similarity 20.2%; Pred. No. 0.62;  
Matches 91; Conservative 80; Mismatches 168; Indels 111; Gaps 25;

Qy 14 ILHEWKFFDYDFGSDERRQDAIILSGEYDYKQNYPSDIDQWHDKIFVTMLRYNGVSPSSLVN 73  
Db 14 VASENSYNDNINTPKEIDTIVS-----NNVENDLDQ-----VILNFS-ESAPGL 60

Qy 74 ISKKVGDGGLLPQVPDWSFAKYDDCSGIVSASKLAIDKCDRLWLDLGLV-----NN--- 126  
Db 61 SDEKLN-----LITQNDAYIPKYDS-NGTSDIEQHDVNLNVPFYLDQAKVPEGENNVNL 114

Qy 127 -----TQPMCSPLTLFDLTTSQLLKQVEIP-----HDVAVNATTCKGRLS----- 167  
Db 115 TSSIDTALLEQPKIYTF--FSSEFINNVNKPQVQALFVSWIQVVLVDFTTEANQKSTVDK 172

Qy 168 ----SLAVQSILDCNTNSDTWVYIADKEGGLIVVHNSDDSFHRLTSN-TFYDYPDKF---T 219  
Db 173 IADISIVVPIVIGIALN-----IGNEAQKG-----NFKDALELLGAGILLEFEBELLIPT 221

Qy 220 KMTIDGESYTAQDQISGMALSPMTNNL-----YSPVASTSLYYVNTPEFTSDYQ 270  
Db 222 ILVFTIKSFGLSSDNKNKVIKAINNALKERDEKWKVEYVSVFVSNWMTKINT-QFNKKEQ 280

Qy 271 QNDIHYEGVQNILDTQSSAKVSKVSGVLFFGLVGDSALGCWNEHRTLERRHIRTVAQSD 330  
Db 281 M-----YQALQNV---NAIKTIIESK-----YNSY-TLEEKNELNKKYDIK 318

Qy 331 TLQMIASMKIKEAPHPVPIFDRIYNR---EYILVLSN--KMQKMNNDNFDD--VNPRI 383  
Db 319 QIENELNKKVSIAMNNI---DRFLTSSISYLMKLINEVKINKUREYDENVKTYLLNVI 375

Qy 384 MNANV---NELILNTRCENPDNDRTPFKIS 410



```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-096-051-4

Query Match      4.6%; Score 99.5; DB 11; Length 2376;
Best Local Similarity 18.8%; Pred. No. 16;
Matches 78; Conservative 63; Mismatches 119; Indels 155; Gaps 18;

Qy 3 LRGSLSKSLPILHEWKFDD-----YDFGSDERRQDAILSGEYDYKNNYPSDID 51
Db 1170 IIRAVSKNKLNSMNFYEVASPTDQELYIFDINGTHQYTVSLVGTGDLNFSYND-- 1227

Qy 52 QWHDKIFVTMLRYNGVSPSSLNVI SKKVGDCGGLPQYPDMWSFAKYDDCSGIVSASKLAID 111
Db 1228 -----NDITAVTDSNGNTLRIRD----- 1246

Qy 112 KCDRLWLVDGLVNNTPCMSPKLLTFLTTSQLLKQVEIPHDVAVNATTG-KGRLSSLA 170
Db 1247 -----PNRMPV-----RVVSPDNQVIMLTIGTNGCLKSMT 1276

Qy 171 VQSILDC-----NTNSDTMVIYADEKGEGLIVYHNSDDSFHRLTSNTP-----DYD 215
Db 1277 AQGLELVFTYHGNSGLLTKSDTGTTFDFYDSEG---RLTNVTFTPTGVVTVNLHGDM 1333

Qy 216 PKFTKMTIDGESYTAQDGSIGSMALSPMTNNLYSPVASTSLYYNTQFRTSDYQNDIH 275
Db 1334 ---KAITVDIESSREEDVS-----ITSNL-----SSIDSFTYMWQDLRNS-YQ---IG 1376

Qy 276 YEGVQNILDTSQSAKVSKSVGLPFLGVLGDSALGCWNEHRTLERHNIR-----TVAQSD 330
Db 1377 YDG-----SLRIIVASGL-----DSHYQTEPHVLGANTANPTVAKNNM 1413

Qy 331 TL-----QMIASMKIK--EAXPHVPIFDRIYINREVILVLSNMQKQVNNDFNDD 378
Db 1414 TLPGENQNLVEMFRKEQAQGVNFGKRLVNGRNLSSVDFDRTTKTEKIYDD 1468

RESULT 13
US-11-096-051-2
; Sequence 2, Application US/11096051
; Publication No. US20050244868A1
; GENERAL INFORMATION:
; APPLICANT: Kekuda, Ramesh
; APPLICANT: MacLachlan, Timothy K
; APPLICANT: Rastelli, Luca
; APPLICANT: Vernet, Corine
; APPLICANT: Ettenberg, Seth
; TITLE OF INVENTION: Ten-M3 Polypeptides and Polynucleotides and their Methods of Use
; FILE REFERENCE: Attorney Docket No. Cura 967
; CURRENT APPLICATION NUMBER: US/11/096,051
; CURRENT FILING DATE: 2005-03-30
; PRIOR APPLICATION NUMBER: 10/038,854
; PRIOR FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 10/455,772
; PRIOR FILING DATE: 2003-06-04
; PRIOR APPLICATION NUMBER: 60/557,978
; PRIOR FILING DATE: 2004-03-30
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 2
; LENGTH: 2715
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-096-051-2

Query Match      4.6%; Score 99.5; DB 11; Length 2715;
Best Local Similarity 18.8%; Pred. No. 19;
Matches 78; Conservative 63; Mismatches 119; Indels 155; Gaps 18;

Qy 3 LRGSLSKSLPILHEWKFDD-----YDFGSDERRQDAILSGEYDYKNNYPSDID 51
Db 1509 IIRAVSKNKLNSMNFYEVASPTDQELYIFDINGTHQYTVSLVGTGDLNFSYND-- 1566

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-525-567-4.rapbn

Qy 52 QWHDKIFVTMLRYNGVSPSSLNVI SKKVGDCGGLPQYPDMWSFAKYDDCSGIVSASKLAID 111
Db 1567 -----NDITAVTDSNGNTLRIRD----- 1585

Qy 112 KCDRLWLVDGLVNNTPCMSPKLLTFLTTSQLLKQVEIPHDVAVNATTG-KGRLSSLA 170
Db 1586 -----PNRMPV-----RVVSPDNQVIMLTIGTNGCLKSMT 1615

Qy 171 VQSILDC-----NTNSDTMVIYADEKGEGLIVYHNSDDSFHRLTSNTP-----DYD 215
Db 1616 AQGLELVFTYHGNSGLLTKSDTGTTFDFYDSEG---RLTNVTFTPTGVVTVNLHGDM 1672

Qy 216 PKFTKMTIDGESYTAQDGSIGSMALSPMTNNLYSPVASTSLYYNTQFRTSDYQNDIH 275
Db 1673 ---KAITVDIESSREEDVS-----ITSNL-----SSIDSFTYMWQDLRNS-YQ---IG 1715

Qy 276 YEGVQNILDTSQSAKVSKSVGLPFLGVLGDSALGCWNEHRTLERHNIR-----TVAQSD 330
Db 1716 YDG-----SLRIIVASGL-----DSHYQTEPHVLGANTANPTVAKNNM 1752

Qy 331 TL-----QMIASMKIK--EAXPHVPIFDRIYINREVILVLSNMQKQVNNDFNDD 378
Db 1753 TLPGENQNLVEMFRKEQAQGVNFGKRLVNGRNLSSVDFDRTTKTEKIYDD 1807

RESULT 14
US-10-194-487-146
; Sequence 146, Application US/10194487
; Publication No. US20060074226A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Deanoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C312
; CURRENT APPLICATION NUMBER: US/10/194,487
; CURRENT FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 146
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-194-487-146
```

